DOCUMENT RESUME

ED 187 115 FL 011 267

AUTHOR Brown, H. Douglas, Ed.; And Others
TITLE On TESOL '77: Teaching and Learning English As a Second Language: Trends in Research and Practice.
INSTITUTION Teachers of English to Speakers of Other Languages.
PUB DATE 77
NOTE 332p.; Selected papers from the TESOL Annual Convention (11th, Miami, FL, April 26-May 1, 1977).
AVAILABLE FROM TESOL, 202 D.C. Transit Building, Georgetown University, Washington, DC 20057 ($10.00).

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Bilingual Education; Black Students; Communicative Competence (Languages); Curriculum Design; Discourse Analysis; Elementary Secondary Education; *English (Second Language); Higher Education; Immersion Programs; Interdisciplinary Approach; Language Laboratories; *Language Research; Language Universals; Listening Skills; Material Development; Reading Processes; Reading Skills; *Second Language Instruction; *Second Language Learning; Spanish; Tagmemic Analysis; *Teaching Methods; Testing; Writing Skills

ABSTRACT

A selection of 27 papers presented at the 1977 TESOL convention is presented. Part one contains the four plenary-session papers which present: a comprehensive view of the teaching-learning process and related interdisciplinary research; the scope of research on language teaching; some of the larger issues in bilingual education; and a perspective on sociocultural variables in the education of black children. Part two, on the theory and practice of teaching, contains eight articles on classroom teaching procedures. Topics concern: a rationale for classroom grammatical explanations, a systematic outline of article usage in English, listening skills, tagmemic theory applied to teaching speaking skills, creative materials development for the language laboratory, reading skills, and understanding the reading process. Part three presents nine articles on recent trends in second language acquisition research. Topics include: an overview of recent trends and models, universal processes and second language data, discourse analyis and English as a Second Language, communicative strategies, and a Spanish immersion program. Part four includes six papers on organizing and evaluating teaching and learning. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED187115

Teaching and Learning English as a Second Language: Trends in Research and Practice

Edited by

H. Douglas Brown
Carlos Alfredo Yorio
Ruth H. Crymes

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY
James E. Alatis
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

FL 011 267

ON TESOL '77

Selected papers from the Eleventh Annual Convention of
Teachers of English to Speakers of Other Languages
Miami, Florida
April 26–May 1, 1977

Teachers of English to Speakers of Other Languages
Washington, D.C.

Copyright © 1977 by the
Teachers of English to Speakers of Other Languages
Washington, D.C., U.S.A.

All rights reserved. Printed in the United States of America

Library of Congress Catalog Card No. 77-99212

Copies available from:

TESOL
455 Nevils Building
Washington, D.C. 20057

Preface

The 1977 TESOL Convention, held in Miami, Florida, April 26 to May 1, marked the beginning of the second decade for TESOL, an organization which now encompasses a host of professional interests and activities related to teaching and learning English as a second language. One of the key emphases which convention chairwoman Joan Morley highlighted for the Miami convention was "balance of communication" *within* interest areas, *across* interest areas, and *outside* specific TESOL interest areas—an appropriate emphasis at a time when TESOL finds itself so rich in its diversity. Yet there is a centrality of purpose in all the interest areas: to seek answers to the question of how persons learn English as a second language and how teachers might better enable learners to meet their goals.

The present volume, a selection of papers presented at the 1977 TESOL Convention, attempts to capture both the unity and diversity of the convention. The unity of the papers in the volume is capsulized in the title which we have chosen: Teaching and Learning English as a Second Language: Trends in Research and Practice. Each paper focuses on either the teaching process or the learning process, and often both. An innovative note is struck in each paper, as indicated by the word "trends" in the title. And almost every paper is practically oriented with solid grounding in research.

The diversity of the TESOL organization is also reflected in the present volume. Topics run a wide gamut of subject matter interest: general state of the art papers, curriculum and course planning, teaching techniques, theory applied to practice, testing, understanding and teaching several skill areas, and research on linguistic, communicative, and affective factors in second language acquisition.

In selecting articles for the volume, the editors have tried to reflect the innovative flavor of research and practice in the profession. A number of the papers, especially the research papers in Part 3, present new concepts in early stages of germination. Other papers add a new perspective, a new focus, or a new understanding to concepts or terms which have become quite familiar.

The 27 papers included here were selected from some 85 papers submitted to us. Even more papers were presented at the convention but for various reasons were not submitted for publication. Of those that were submitted, a number of papers, by mutual agreement of the editors, were selected for possible inclusion in the *TESOL Quarterly* or some other TESOL publication. So, the present 27 papers are a small fraction of the papers presented at the convention; nevertheless, we feel that this selection is representative and indicative of the subject matter of the 1977 TESOL Convention. Since all of the convention papers were orally presented, many of the papers printed here preserve the informal style of the oral presentations.

Part I, Keynotes, contains the four plenary-session papers which were submitted to the editors. It was our decision to print all of these specially invited papers in one section. All four papers take a somewhat broad overview of a particular topic. Pit Corder gives us a comprehensive view of the teaching-learning process with an implicit link to related interdisciplinary research. Richard Tucker refers more explicitly to a tremendous scope of research in directing his comments to the question: can language be taught? Hernan LaFontaine speaks to some of the larger issues in bilingual education and Donald Henderson provides a unique perspective on sociocultural variables in the education of black children.

Part 2, Teaching: Theory and Practice. consists of eight articles on classroom teaching procedures, all carefully based on linguistic and psychological research. The subject matter is not comprehensive, but is a representative sample of important topics in theory and practice: a rationale for classroom grammatical explanations (Patricia Furey), a systematic outline of article usage in English (William Acton), listening skills (Anna Uhl Chamot), tagmemic theory applied to teaching speaking skills (William Crawford and Beverly Fried), creative materials development for the language laboratory (Wu Yi So and Doris Scarlett), and three articles relating to reading skills and understanding the reading process (Pat Rigg, John Schafer, and Mark Clarke and Linda Burdell).

Part 3, Learning: Research in Second Language Acquisition, presents nine articles on recent trends and experimental studies in second language acquisition research. The past five or six years has seen a mushrooming of interest and inquiry into the acquisition process. The papers in this section are representative of practical research which has direct relevance to teaching: an overview of recent trends and models (Stephen Krashen), universal processes and second language data (Georgette Ioup and Anna Kruse), two articles on discourse analysis and ESL (Diane Larsen-Freeman and Sue Vander Brook, Karen Schlue, and Cherry Campbell), two studies of communication and communicative strategies (Elaine Tarone and Stephen Gaies), some data from children in a Spanish immersion curriculum (Sandra Plann), the study of affective variables, in this case, self-esteem (Adelaide Heyde), and another affectively oriented article, an introspective, longitudinal study of language acquisition (Francine Schumann and John Schumann).

Part 4, Organizing and Evaluating Teaching and Learning, combines aspects of both teaching and learning through such topics as developing curricula, understanding classroom processes, and testing and evaluating language competence. The six papers particularly examine: curriculum and syllabus organization (Bernard Mohan and William Rutherford), a new definition of some misunderstood terms in teaching/learning (Peter Strevens), a new perspective on how teachers can give feedback to classroom learners (Michael Long), proficiency testing (Donald Bowen), and the always difficult task of assessing writing proficiency (Karen Mullen).

In the making of this volume, the editors are deeply grateful to Joan Morley, chairwoman of the 1977 TESOL Convention, whose leadership and efficiency in planning the convention were appreciated by so many. In the words of the resolution passed by the TESOL Legislative Assembly on April 29, 1977, "Joan Morley, with high communicative competence and outstanding organizational performance, has guided us to a convention program that is unprecedented in its professional contribution—in depth, in breadth, and in timely significance." Because of a well-deserved leave of absence, she asked to be relieved of the responsibility, normally given to convention chairmen, of editing the proceedings of the convention. Her name is therefore not listed as a co-editor; however, the excellence of the papers which were presented at the convention are due in large part to the wisdom of her planning.

Others also helped to bring the volume into existence: John Upshur gave us invaluable comments on some of the papers submitted to us; Katherine Murphey has given countless hours of secretarial and clerical work; and the interest and support of James Alatis and Carol LeClair in the TESOL office have been greatly appreciated.

We hope that *Teaching and Learning English as a Second Language: Trends in Research and Practice* will not be just a sentimental memoir of the 1977 TESOL Convention, but also a comprehensive representation of the state of the art of TESOL in 1977, a volume which will be useful as a textbook for teacher trainees and as a handbook for professionals and paraprofessionals in the field of teaching English as a second language.

—H.D.B.
C.A.Y.
R.H.C.

Table of Contents

TEACHING AND LEARNING ENGLISH AS A SECOND LANGUAGE: TRENDS IN RESEARCH AND PRACTICE

ON TESOL '77

Preface iii

Part 1. *Keynotes*

Corder, Pit: Language Teaching and Learning: A Social Encounter . 1
Tucker, G. Richard: Can A Second Language Be Taught? 14
La Fontaine, Hernan: Bilingual Education and ESL: A Symbiotic Relationship 31
Henderson, Donald M.: Cross-Cultural Interference in Black-White Relations: Discord in the Classroom 36

Part 2. *Teaching: Theory and Practice*

Furey, Patricia R.: The Preparation of Grammar Explanations in the Teaching of English As A Second Language 46
Acton, William: A Conceptual Framework For Teaching Articles In English 57
Chamot, Ana Uhl: Listening Skills in the Elementary ESL Class . . 74
Crawford, William and Fried, Beverly: Tagmemic Concepts in the ESL Speaking/Pronunciation Class 81
So, Wu Yi and Scarlett, Doris L.: A New Direction for Language Lab Programs 94
Rigg, Pat: The Miscue-ESL Project 106
Schafer, John C.: Lessons Learned from Reading a Vietnamese Text: The Importance of Pre-texts in Second Language Reading . . 119
Clarke, Mark A. and Burdell, Linda: Shades of Meaning: Syntactic and Semantic Parameters of Cloze Test Responses 131

Part 3. *Learning: Research in Second Language Acquisition*

Krashen, Stephen D.: Some Issues Relating to the Monitor Model . . 144
Ioup, Georgette and Kruse, Anna: Interference versus Structural Complexity in Second Language Acquisition: Language Universals as a Basis for Natural Sequencing 159
Larsen-Freeman, Diane: A Rationale for Discourse Analysis in Second Language Acquisition Research 172
Vander Brook, Sue, Schlue, Karen and Campbell, Cherry: Discourse and Second Language Acquisition of yes/no Questions 178
Tarone, Elaine E.: Conscious Communication Strategies in Interlanguage: A Progress Report 194

Gaies, Stephen J.: The Nature of Linguistic Input in Formal Second Language Learning: Linguistic and Communicative Strategies in ESL Teachers' Classroom Language 204
Plann, Sandra: Acquiring a Second Language in an Immersion Classroom 213
Heyde, Adelaide W.: The Relationship Between Self Esteem and the Oral Production of a Second Language 226
Schumann, Francine M. and Schumann, John H.: Diary of a Language Learner: An Introspective Study of Second Language Learning . 241

Part 4. *Organizing and Evaluating Teaching and Learning*

Mohan, Bernard A.: Toward A Situational Curriculum 250
Rutherford, William E.: Parameters of Language Syllabus Construction 258
Strevens, Peter: Causes of Failure and Conditions For Success in the Learning and Teaching of Foreign Languages 266
Long, Michael H.: Teacher Feedback on Learner Error: Mapping Cognitions 278
Bowen, J. Donald: Practice Effect in English Proficiency Testing . . 295
Mullen, Karen A.: Using Rater Judgments in the Evaluation of Writing Proficiency for Non-Native Speakers of English . . . 309

9

8

Part 1

Keynotes

Language Teaching And Learning: A Social Encounter

Pit Corder

University of Edinburgh

The language teaching classroom is the locus of a special type of social encounter. Teacher and learner each bring to the classroom certain expectations of each other's behaviour, a definition of each other's roles. If this encounter is to be fruitful, if learning is to be both relevant and efficient, there must be mutual agreement about what constitutes appropriate behaviour. In part these matters are culturally determined but in part negotiated in the classroom itself. Part of language learning is discovering the language learner's role.

But there are in addition certain minimum requirements of teacher and learner behaviour which are culture-and classroom-independent. These are the 'linguistic' aspects of language learning and teaching. In what way are these 'universal' requirements modified by culture-specific expectations? In other words how do good 'general pedagogy' and good 'linguistic pedagogy' interact? The problem of reconciling the particular culture-and classroom-bound on the one hand with the universal on the other is not peculiar to language teaching alone. It is found in the teaching of any school subject. But what makes language teaching a special case is not that it alone is a social encounter but that it alone is a *social* encounter to teach *social* behaviour.

We have heard a great deal recently about native speaker-hearers of languages, ideal or otherwise, but it is an interesting thing that, while we find nothing odd about enquiring whether someone *speaks* French or German, say, we do not ask whether someone *hears* French or German. (Though we do, of course, enquire whether someone understands French or German). It seems at first sight as if we can't enquire as to performative ability in a language without specifying one or the other role, and yet we have words in English which refer to language activity which are neutral in this respect—such as *converse*, *commune*, *discourse* or simply *talk*, or in the written mode *correspond*. It is curious that we never ask whether someone can *talk* in English or *converse* in French.

And yet when we want only to indicate that language activity is going on we happily say of some group of people that *they are talking* or that *they are conversing in French*, as we might say that they *are discussing their plans* or *they are arguing about politics*, thereby focussing attention on their joint cooperative activity—on the interaction of the participants—on their interlocking role-behaviour.

When we turn to the institutionalised activity of developing in someone the ability to *talk*, *converse*, *commune* or *correspond* in a language, English, most unfortunately, obliges us to focus separately on the roles of the participants in this activity—we must either speak of *teaching* or *learning*—there is no single word which can be used to refer to this activity in the neutral way that *talk* or *converse* can refer to the activity of verbal interaction. That is why my title today reads, infelicitously, in part: *Language teaching and learning.* Perhaps we could usefully invent the word "teach-learn" and thus refer to what may go on in a classroom, for example, as "teach-learning" and in reply to the question: "*what are they doing in there?*" reply "*Oh, they're teach-learning English*", as one might say "*Oh, they're talking English*" or "*discussing philosophy*".

The object of this short excursion into the semantics of English is to draw attention to the Whorfian effect that our language has on the consideration of the language learning process. It has made it difficult for us to view it as a cooperative enterprise and has forced us into the position of treating it as two separate, potentially unco-ordinated or unconnected activities—those of teaching and learning. We all too readily write articles on "teaching method" or institute research programmes into "language learning" and are happy to assent to the proposition that we can learn a language without a teacher, and even, though perhaps less happily, that someone can conscientiously teach something without his pupils really learning anything.

My thesis today is that language acquisition, whether of the mother tongue or a second language, is a cooperative enterprise, and that it takes place uniquely through a process of social verbal interaction or *talk*.

You may object at this point that there are cases of people who have taught themselves a language. The easy but incorrect answer to this objection is that they have then combined in themselves both the role of teacher and learner, as one might argue that one who talks to himself combines both the role of speaker and hearer in one person. I would suggest rather that when we examine such cases we find that they have typically made use of either so-called self-instructional materials *or* ordinary written or spoken texts in conjunction with a dictionary and perhaps a grammar book. But let us remember that these materials are themselves samples of *discourse* between a writer and a reader, and furthermore, that self-instructional materials, dictionaries and grammar books are examples of specifically pedagogical discourse.

I will not even bother to consider the case of someone who is said to acquire a language by living in the community where it is spoken or by interacting with speakers of that language, whether native or not. These cannot, on any reasonable interpretation of what is going on, be examples of self-instruction. I would, on the contrary, maintain that the people with whom such learners interact can properly be said to be teaching, since they typically manifest teacher-like behaviour by modifying their speech in a variety of ways which make it easier for the learner to understand their talk and thus learn the language.

However, while we can say that such interlocutors "teach", we would not wish to call them "*language* teachers" and it is with institutionalised "language teach-learning" that I am concerned today. In this situation we have to do with an interaction between participants in the roles of pedagogue and pupil as they may be institutionalised in different cultures. Now, as we know, the justification for analysing social life in any culture in terms of the roles people play is that this focuses attention on the social bargains which underlie cooperation. Every social relationship consists of a pair of roles and when someone enters voluntarily one of these roles (and I emphasise voluntarily) he does so because he believes the relationship will help him to further some of his objectives (whether these can be regarded as altruistic or not). Every member of some social institution, and in our case we are concerned with educational institutions, has a part to play. He has tasks to perform and is entitled to receive services from others in recognition of his contribution. These bundles of rights and obligations are what we mean by roles. And these rights and obligations are realised in the socially sanctioned behaviour of the participants in the cooperative activity.

The first thing we have to note, therefore, is that the language teacher and language learner in any culture operate within the range of behavioural freedom accorded by that society to the roles of *pedagogue* and *pupil*. If you like, a *language teacher* is a *pedagogue* first and a language teacher second, and the *language learner* is a *pupil* first and a language learner second. By this I mean that a language teacher or language learner cannot normally together indulge in behaviour which is not otherwise sanctioned in the pedagogue-pupil relationship in that society. But in most cultures, both language teachers and language learners bring with them to their interaction a number of *specific* expectations of each other's behaviour as to how the "teach-learning" of language is to proceed. Thus, for example, it is quite widely expected by language learners that the teacher will prescribe a particular text-book and use it; or that work in the language laboratory will be part of the learning activity, or that drills will be indulged in and, more generally, that it is the teacher who directs or controls the interaction. One could easily enumerate for any particular educational institution or system in a culture the expectations that both language teacher and language learner bring to the classroom about what constitutes acceptable and unacceptable "language teach-learning" behaviour.

One would, of course, not wish to suggest that institutions in a society are necessarily static, or that roles within a society are not liable to undergo continual redefinition or renegogiation; nor, and I think this important, that new educational theorising does not sometimes diffuse (slowly perhaps) throughout society, or parts of society, in such a way that a redefinition of acceptable language "teach-learning" behaviour results, thus allowing prevailing notions about *good* language "teach-learning" to be realised—so long as these do not extend beyond the current norms of the sanctioned roles of pedagogue and pupil. Many of us are surely personally aware that from time to time we have wished or attempted to try out "teach-learning" activities in the classroom, only to be frustrated by the school or college authorities, or even the pupils themselves.

Now clearly there are great differences in the degree of freedom accorded by different societies to the roles of pedagogue and pupil. In some cultures we know of, we may feel that these roles are so constrained that *good* language "teach-learning", as we conceive of it, cannot take place, and no negotiation is possible which can alleviate the situation. In other societies the roles of pedagogue and pupil are so loosely defined as to permit a great deal of negotiation within the classroom as to the nature of the interaction and the consequent language "teach-learning" activity.

The question that poses itself at this point is whether it is ever the case that the roles of pedagogue and pupil are so constrained that language "teach-learning" is effectively inhibited, or actually prevented. Or, to put it another way, does language learning require some minimal conditions to obtain such that, if they do not, it cannot take place? It is to this latter question that I shall shortly address myself; but before I can do so it is necessary to say something about the goals of the interaction which I have called language "teach-learning".

You will recall that the notion of role was connected with the social bargains which underlie *cooperation*. Clearly "teach-learning" is a goal-oriented interaction in which both participants agree that the object for which they are *cooperating* is the achievement of a so-called "knowledge of the target language". We can say that one element in the expectations that each party brings to the other's role is that they share a common goal. What is, however, not automatically the case is that both will agree upon just what is meant by a "knowledge of a language". That is why it is necessary to say something about this before going on to deal with the question I have already posed, that is whether there may be such a constraint on the roles of pedagogue and pupil in some cultures that an achievement of this goal is effectively inhibited or prevented. It is also necessary to point out here that, in recent years, research and theorising into the nature of human language by linguists may lead us to redefine what we mean by "functional knowledge of a language" in such a way as to cause us to give a positive answer to this question. Indeed I am at present reluctantly coming to the conclusion that our more recent notions about what constitutes a "functional knowledge of a language" are such that perhaps all classroom settings are so constraining that really efficient language teach-learning cannot take place in them at all. This is a notion to which I shall return later.

So what might we mean by "functional knowledge of a language" and to what extent do we find agreement between the participants about the goals of the "teach-learning" interaction? I do not wish to dwell long on this topic. It is, I think, generally, but somewhat vaguely, agreed that the goal of any language "teach-learning" is to enable the learner to interact with speakers of a language, native or otherwise, for some greater or smaller range of cooperative purposes, that is, engage in *talk* or *correspondence*. I emphasize that it does not necessarily have to be native speakers who are involved. There is still a totally mistaken belief that a native-like *accuracy* in the production of that range of linguistic forms thought appropriate for the learners' purpose is the only thinkable level of achievement. There are, however, innumerable situations of verbal interaction where a far lesser achievement is a perfectly adequate functional objective for particular groups of learners. Some grammar of a pidgin-level of structural complexity, for example, is for many learners of a language an entirely appropriate goal for their social talking needs, the achievement of which would constitute success in their terms. As far as the knowledge of the formal properties of a language is concerned, it is not the accurate productive mastery of some *limited range* of standard language structures which is the appropriate objective but the confident possession of a knowledge of a structurally less complex grammatical system. To suggest this to most language teachers at the present time is usually considered as rank heresy.

But what is important is not just that the learner should have available some grammar or other and some vocabulary sufficient for his talking needs but that he should also command an appropriate and effective rhetoric—that is, the ability to *use* appropriately whatever grammar he has to satisfy his needs as a talker. Now we are, of course, all aware that language "teach-learning" in the past (and still largely at the present) has had as its principal goal a working knowledge of the structural rules of the standard language, that is, the fully-complex target language code, and that the bargain between language teacher and language learner, which underlay their cooperation, was for the achievement of this objective. Both agreed, if you like, upon what we now see as an inadequate definition of what is meant by "functional knowledge of a language". Neither envisaged language as *social behaviour*; neither conceptualised language as *doing something*. Both thought of it rather as "knowing something". It was natural, therefore, that "teach-learning" a language should have resembled "teach-learning" other so-called "content subjects" in schools or colleges.

I will immediately counter the objections you will raise to this analysis, that, on the contrary, teachers at least have long thought of language learning as the acquisition of a body of skills. I would suggest, however, that most of the classroom activities which might be put forward as evidence for this—such as drills, reading and writing practice and so on—are in fact (when not merely testing procedures) practice in making the structural knowledge already acquired efficiently and rapidly available for producing and interpreting well-formed sentences or, (more recently) well-formed texts—not, be it noted, coherent discourse or talk. The distinction is important. Sentences and texts are abstract linguistic objects; utterances and discourse are sentences and texts put to work.

I can now attempt to answer the question I posed: Does language learning require some minimal set of conditions to obtain such that if they do not, it cannot take place? The short answer is that if we adopt the restricted analysis of "knowledge of a language" as the ability to produce or process well-formed sentences or texts, that is, a knowledge of the grammatical or structural properties of the language, then there are unlikely to be societies or cultures where the constraints on the roles of *pedagogue* and *pupil* are such as to prevent what has to be done to achieve such knowledge from being done. On the other hand, if our definition of language as "*doing something together*" or "*talk*" is accepted, then it may well be that there are indeed few educational institutions where the constraints on the roles of pedagogue and pupil are so loose as to permit all the possible goals of language "teach-learning" to be realised. In other words, no classrooms anywhere permit all possible language learning goals to be achieved.

Now you may feel that this is a trite observation. Of course, we are not able, and indeed never set out, to achieve *in full* the goal that every learner should be able, at the end of his course, to walk out of the classroom a fully-fledged talker or corresponder for his purposes in the target language. All we can do is to give him the tools, the possession of which will enable him to finish the job himself—help him, if you wish, to develop a set of efficient language learning strategies.

Even if we concede this point, there still remains the question as to what the minimum conditions are that must hold for us even to achieve this more limited goal.

I have so far made a distinction between the roles of *pupil* and *pedagogue* as defined in any society and those of *language teacher* and *language learner* and suggested that the latter pair are specialised or particular cases of the former—that is, the professional language teacher has both what I shall call the *general* role of pedagogue and also the *technical role* of language teacher, whilst the learner has the *general role of pupil* and the *technical role* of language learner. The analogy is with the general role of soldier and the particular role of, say, sergeant. Let us now consider the behaviour that is associated with their technical roles. I am adopting for this analysis the theoretical point of view that language learning is an inductive cognitive process of discovering those regularities which underlie talk in the target language, which are often called the rules of grammar and the rules of rhetoric and interactional management, or the rules of *usage* and the rules of *use* respectively. (Not all linguists, of course, would wish to make a sharp or indeed any distinction between these; c.f. Halliday). The processes involved are basically those of data processing, hypothesis formation and testing. Clearly the first requirement for learning to take place is that the learner shall be exposed to, (or have available), the data upon which these processes can operate. Note here that if the goal is defined only as a knowledge of the structural properties of the target language, then the data is viewed as sentences or texts in the language. If, on the other hand, the goal of learning is defined as the ability to *talk to some purpose* then the data is viewed as samples of discourse or talk. Or to put it another way—if the learner is exposed only to linguistic forms, that is sentences or texts, in the target language all he can discover will be the rules of the grammar or the structural properties of the language code (rules of usage), and all he will be able to do at the end is produce more or less well-formed sentences or texts.

In order to learn to talk or correspond he must be exposed to *discourse* in the target language, that is, authentic examples of people talking or corresponding. Only thus can he discover the rules of use, or rhetoric. His ability to participate outside the classroom in creating the many types of discourse which it may be his wish or need to do, will be related to his experience of these during his learning career. Just as we cannot expect him to discover rules of structure which have not been in one way or another exemplified in the sentences he has been exposed to, we cannot expect him to discover the rules of use or rhetoric which have not in one way or another been exemplified in the discourse he has been exposed to.

It is then part of the expected technical role of the language teacher that he should make available the data on which the learner is to operate. But any competent informant can do this. For an informant to qualify as a teacher, however, it is conventionally understood that the data he provides is controlled, adapted, selected or otherwise organised in such a way as to make the processing of it by the learner easier. I have already pointed out that some sort of simplification of the grammar or the rhetoric is practised by a mother or native-speaker when interacting with learners, either of the mother tongue or second language, even though these cannot be called professional language teachers. I suggested, nevertheless, that their behaviour could properly be regarded as teaching precisely for this reason. Provision of data in some sort of processed form is one part of the generally agreed technical role behaviour of the language teacher.

If the goal of "teach-learning" is no more than a control of the grammatical rules of the language, then the name "informant" is an appropriate one for the teacher, connected as it is with the notion of *information*. The teacher *has* the information and he *gives* it to the learner. I suggest we classify teachers who perceive this as their principal task as acting technically in the role of "informants" or "knowers", that is, possessors or sources of knowledge about the language which the learner is trying to get at. The learner appears in the role of "information-seeker". Note that the relationship implied by this analysis is one which necessarily places the learner in a totally dependent position. This is a role relationship of pedagogue and pupil with which we are all too familiar, alas, in rigid and authoritarian educational institutions.

Access to processed data, however, is probably not by itself a sufficient condition for language learning to take place. If this involves hypothesis-formation and testing, then the learner must have the opportunity of submitting his hypotheses to confirmation, or otherwise, by producing sentences or texts for assessment by his "informant". This confirmation or otherwise of his hypotheses by the "informant" or "knower" is part of the technical role of the language teacher. Disconfirmation takes the form of correction; confirmation, of acceptance. This, of course, may be done in many ways from overt correction, rephrasing, praise, smiles, nods or refusal to understand. Again any "informant" who has some knowledge of the language, native or otherwise, can and does behave in this way. He is, for this reason I think, properly said to be teaching though he is not acting in the role of pedagogue. One may add here that the amount and nature of this teaching behaviour permitted outside the classroom is probably culturally determined. In many societies it is socially unacceptable to perform correction except in a highly mitigated form. Even rephrasing is frequently interpreted as a form of rebuke.

If the participants in "teach-learning" agree on a knowledge of the structural rules of the language as the only goal of their interaction, then, I believe, the provision of the language data, the elicitation of hypothesis-testing behaviour and the provision of feed-back are all that are required of the technical role of the language teacher. Anything over and above these is a matter of negotiation between the participants, or a sanctioned part of the general role of pedagogue. There is, of course, much beyond this that is typically observed in language classrooms, notably all that "teach-learning" behaviour which can be called guidance, prompting, cuing, explaining, describing and so on; that is to say, all that teacher-behaviour aimed at helping the learner to discover more quickly and efficiently the formal properties of the language system. If we analyse the discourse created by "teach-learning" in the language classroom we will find that the major part of it is represented by exchanges which consist of cycles of presentation, testing and feedback.

The picture I have drawn is no doubt idealised and simplified; much other discourse no doubt is created in the language classroom. But what I would suggest is that if we examine it, it is not part of the "technical behaviour" or discourse of the "language teach-learning", but derives from the more general relationship of pedagogue-pupil; for example, all that talk which is concerned with organisation, negotiation, planning and discipline.

I suppose, however, most of us here today regard "language teach-learning" as having the more comprehensive goals I spoke of earlier, the enabling of the learner to function as a *talker* or *corresponder* in the target language in some limited set of social encounters. This goal corresponds to a view of language as social behaviour, as *doing* rather than *knowing*, and involve the discovery of the rhetorical rules of the language as it is used in some particular social group or sub-group within the language community. What I want to suggest now is that, if we negotiate an agreement with our pupils that this is to be the goal of our cooperation, then it necessarily implies a redefinition of the *technical* roles of language teacher and learner.

It is at this point that a conflict, or mismatch, may arise between the behaviour of pedagogue and pupil which is sanctioned in some educational institutions or cultures and the necessary technical behaviour of language teacher and language learner. In other words, if we agree to aim at this goal then the norms of the educational system may have to be broken or extended if we are to achieve it. I believe this to be a conflict that has not yet been wholly resolved anywhere, since nowhere do the conditions exist in educational institutions for it to be so. It is this that led me earlier on to say that the language classroom may perhaps not be the best place to "teach-learn" languages in.

In many places the first stages towards a redefinition of the technical roles of teacher and learner have already been taken, and attempts are being made to "teach-learn" so-called "communicative competence" in the classroom. Perhaps redefinition in this case has been more in the nature of an extension of the role repertoires so far described. What has happened is that the teacher has added to his role of informant or "knower" that of "producer" and the learner to his role of "information-seeker" that of "actor". This extension of the role behaviour of the participants now leads to a more complex network or pattern of interaction in the classroom. The learner is no longer just attempting to discover the regularities underlying the sentences or texts *of* the language but to discover the social norms of some types of talk *in* the language. And he is doing this not just by interacting with the teacher but through interaction with fellow learners in simulated speech events, in role-playing exercises.

This drama is now being monitored by the teacher as a producer monitors the rehearsal of a play. I think it important, however, to insist at this point that this is only *playing at talk*, not *talking*, just as, for example, practice in writing letters, essays and other compositions for the teacher is *playing at correspondence*, not corresponding. It is "learning by playing" and crucially depends upon the interactional management skills and rhetorical knowledge of the teacher for its success. Where the teachers are native speakers of the target language, or more exactly, members of the social group whose rhetoric is being taught, then some measure of success may be achieved; but the great majority of teachers of languages are not members of the community and culture of the language they teach, and this goes, of course, for English teachers too. In these cases one has observed a notable reluctance by teachers to adopt in the classroom the role of producer and by learners to adopt the role of actor. This is in part because "learning by playing" is often not socially sanctioned behaviour in the classroom anyway; but probably equally because the teacher feels insecure in his command of the rhetoric of any English-speaking community and consequently not competent to monitor critically the talk of his pupils.

We can look upon role-playing exercises as a cooperative problem-solving activity by the learners. If we do this, however, we have to ask ourselves what the nature of the problem that the learners are attempting to solve actually is. Supposing the play is about buying an airline ticket and the learners take the parts of counter clerk and customer, in what sense can they be said at the end of the simulated interaction to have solved a problem? The only problem it seems that they may succeed in solving is satisfying the teacher that their discourse has verisimilitude. Now I do not wish to underestimate the value of this sort of "teach-learn" activity. Clearly acceptance by the teacher, where granted, must indicate that the learners have discovered some of the rules of rhetoric and transactional management—though just where they have induced them from is a matter of some interest, to which I shall return. It is clear, however, that if we regard any goal-oriented interaction *outside* the classroom, such as buying an airline ticket, as a problem-solving or goal-oriented type of interaction, then success is demonstrated by the customer getting the ticket he wants and the airline its money, that is, the legal performance of a commercial transaction. There is a fundamental difference in the nature of the authentic transaction and the simulated one, namely, the motivations and intentions of the participants. This is precisely the relation between the theatre and real life.

There are, however, problem-solving exercises in the classroom which are authentic. These are of the sort where teacher and pupils set themselves a problem in which success is demonstrated by the *outcome* not by the *process*—that is, in which a solution is discovered *by means of talk*, for example, the learners attempt to find the answer to some factual question, sort pictures into a logical sequence, construct a model from pieces, play a game and so on. These are authentic because all of them are cooperative activities which are also regularly found outside the classroom. Exercises like these are often highly valued pedagogically because they represent what is sometimes called "using the language for genuine communicational purposes". The only comment I would make here is "what language?" When the *outcome* is in focus the *means* will probably not be. This, I may say, doesn't worry me, but it will worry a lot of teachers who hold to the "all-or-nothing" view of "language teach-learning" I spoke of earlier. Such teachers will be unwilling to use group-work problem-solving exercises, not just because they may not be sanctioned within the educational institution, but more likely because they cannot effectively monitor the *formal* aspects of the language which occurs in these interactions. This attitude betrays an adherence to the old notion that it is better to keep silent than to practice incorrect forms, and, ultimately, that *means* are more important than *ends*. My own belief is that anything that promotes effective "strategies of goal-directed talk" is to be encouraged. Part of the learning of the role of language learner which I spoke of before is the acquisition of such strategies.

Teachers who promote such "teach-learning" activities have added another behaviour to their technical role repertoire of "informant" and "producer". Shall we call it "referee"? And the learner too has added to his repertoire that of "player".

I think it is worth noting at this point that as each new extension to the behaviour repertoire of the participants in the "teach-learning" activity is made, the role relationship of pupil and pedagogue undergoes change. The learner is becoming increasingly responsible for the conduct of his own learning and thus in a less dependent relationship to the teacher. Teacher and learner are becoming more equal *partners* in a cooperative enterprise.

The question is what is the next step? What I have described so far is what *has* happened, and *is* happening in those educational institutions within which the sanctions operating on pedagogue and pupil do not inhibit such behaviour. The next step is the biggest yet to be recorded, but it can only happen where the target language is that of the community in which the institution is situated—that is, in a so-called "second language situation". This is because it involves him in leaving the classroom and going out into the community. When this happens the participants in the "teach-learning" activity are not just teacher and learner, but native-speakers in society outside the classroom in their various everyday roles. They may, of course, come into the classroom, but the range of discourse which is created by this is generally limited to lecturing or discussion. What I have in mind is where learners, with or without the teacher, go out into the life of the community to interact with members of the community in the relationship of customer to shopkeeper, patient to doctor, employee to employer, lodger to landlord, guest to host. Now you may say that, if this is indeed a "second language learning situation" this will happen without the intervention of the teacher. This is perfectly true. But what I have in mind is that these encounters should not simply be fortuitous, but systematically organised and negotiated by the teacher and learners with the members of the community, who are drawn in as an integral and planned part of the "teach-learning" process. Learners are able not only to observe authentic talk of many different, and for them relevent, types but also participate, after classroom simulation, perhaps, in the roles I have exemplified, and again, after the encounters, discuss authentically and purposively in the classroom their experiences with each other and the teacher. This is *learning by doing*. The teacher has now added perhaps the last behaviour to his repertoire, that of "mentor" and the learner that of "apprentice member of society".

What I have tried to do so far in this paper is to treat language "teach-learning" as a social encounter of a particular sort. What is produced in social encounters is *talk*. Language "teach-learning" is thus a particular sort of talk or discourse. Whenever people interact the sort of discourse that ensues is determined by the relationships of the participants and the agreed goals of their interaction. The discourse produced in any interaction is the means whereby the goal of the participants may be negotiated and achieved. This is, of course, true of any "teach-learning" interaction, not just of language "teach-learning". But what distinguishes language "teach-learning" is that it is, in a sense, *reflexive*. By this I mean that the ability to talk or correspond is achieved through talk or correspondence. Thus, talk or correspondence is both the end and the means. It is a perfect example of learning by doing. What determines in any particular encounter the nature of the talk that is generated is the definition that the participants bring to the encounter of what constitutes an ability to talk or correspond. There is, herefore, an interdependence between goals and role-relationships. If the definitions which the participants bring of the nature of this ability do not match, then there will be a conflict as to their definitions of their respective role-relationships in the interaction, and their goals will not be achieved. In successful language "teach-learning", therefore, there must be agreement between the participants about the nature of their goals, that is, about their analysis of what constitutes the ability to talk and correspond in the target language. It is on this basis that their respective roles must be negotiated. Failure to agree upon goals will lead to some participants simply opting out of the encounter.

If, for example, the teacher perceives his role as "knower" but the learner as "actor" or "apprentice member of society" then this implies a mismatch of definitions about their respective role-relations and successful "teach-learning" will not take place, or if the learner defines his role as "information seeker" and the teacher as "mentor", again successful "teach-learning" will not take place. Conflicts of goal definition and respective roles show themselves, I suggest, in the generally massive failure to teach foreign languages successfully in school systems where language learning is compulsory. It also often shows itself where attendance is voluntary, in the frequent rapid fall-off in attendance in the early stages of a course.

I believe we shall not be successful in language "teach-learning" so long as we insist on the traditional relationship of pedagogue and pupil in which the teacher is seen as the dominant member in this relationship, the hander-out of information. The status of teacher in this situation was based on the belief that the teacher possessed something which the learner wanted and did not possess. The teacher possessed skills and knowledge which the pupil did not. Now this may be true of the teacher of physics or history, but it is not, I believe, and never was, fundamentally true of the language teacher. So long as the definition of the learning goals of language "teach-learning" was a "knowledge of the structural features of the target language", then the teacher was indeed in possession of something which the learner did not have; but when we recognise that becoming an affective talker in a language involves a knowledge of the rhetoric (or rules of use) of a language and interactional management skills current in a particular language community, then the teacher is in a weaker position. Firstly, he does not have an *explicit* descriptive knowledge of these rules, even if he is a native speaker, certainly not in the sense that he has an explicit descriptive knowledge of at last some of the structural properties of the language. Consequently he cannot teach these systematically. Secondly, the learner, by virtue of being a competent talker or correspondent in his mother tongue, already possesses *a* rhetoric which will serve, at least initially, for talk in the second language, even if the eventual target rhetoric is not identical with that of his mother tongue. This is the answer to the question I raised earlier: where does the learner get his rhetoric from in role-playing exercises?

But there is now another reason for considering that the teacher does not possess what the learner needs. Although he possesses, as I have said, an explicit knowledge of some of the structural properties of the target language, it is now becoming probable that this knowledge may not be as important as we used to think. The dominant position held by the teacher in his role of "knower" was based on the belief that whatever the teacher decided to teach was in principle learnable by the learner at any time. Thus, it was the teacher who knew what was to be learned and the sequence in which it was to be acquired; it was the teacher who determined the way in which the learner was gradually to elaborate his structural knowledge of the language. It now appears that the process may not be as simple as that, and indeed that it is the learner who *takes* the data and processes it when he *needs* it and can assimilate it. It is the learner who calls the "structural tune" and not the teacher, just as it now appears that it is the infant learning his mother tongue who determines the course that the elaboration of his linguistic knowledge takes and who controls his mother's verbal behaviour. It is the mother who responds to the child's initiatives and needs, and not vice versa. "Good" or "appropriate" teaching is, therefore, perhaps no longer to be seen as imposing a highly organised and detailed syllabus upon a group of learners and as a process of *putting in* or *handing out* information, but as the task of responding to the developing functional or talking needs of the learner by making the appropriate data for learners available "on request". Good language "teach-learning" discourse is not teacher-directed but learner-directed.

The problem of realising good "teach-learning" is that we still largely work within the framework of educational institutions which impose an interpretation of the pedagogue-pupil relationship which accords the teacher a dominant status and expects of him that he will direct the "teach-learning" programme. There are, of course, situations already where this is changing and already named "methods" exist which propose a different relationship and consequently different participant behaviours (the Silent Way, Community Language Learning, Individualised Instruction, etc.). I, personally, am somewhat suspicious of all proposals for "new methods". The history of language teaching is punctuated by a series of named methods, each of which was hailed in its day as the answer to all the problems of language "teach-learning". In the end successful language "teach-learning" is going to be dependent upon the willing cooperation of the participants in the interaction and an agreement between them as to the goals of their interaction. Cooperation cannot be imposed but must be negotiated. And what can be negotiated will depend, in the end, upon what the participants accept as negotiable—that is, upon what society sanctions as acceptable "teach-learning" behaviour. I believe that the changes one can observe at the present time in educational thought and practice, at least in western industrial society, is moving in the direction that will make the last scenarios I have drawn increasingly likely. We are seeing a general movement towards students demanding relevance in their studies, and pedagogues instituting discovery methods of study in many fields. The student of the future is not going to unquestioningly accept that the "teacher knows best", but is going to demand of him a justification, a reasoned argument for his programme. And that argument is going to have to be based upon a more adequate account than is presently current of how one becomes a competent talker in a language.

Can A Second Language Be Taught?

G. Richard Tucker

McGill University

During the past decade there has occurred a great deal of theorizing and research about various facets of second-language learning and teaching. Much of the work in North America has been prompted by the challenge to provide effective second-language teaching within the context of the public school system, while simultaneously nurturing the native language development and sociocultural traditions of heterogeneous student populations. In many other countries, decisions have recently been taken to decrease the use of a foreign language as a medium of instruction or even to decrease the amount of time devoted to teaching the language as a subject despite the fact that many citizens in these countries retain a demonstrated need for greater foreign-language proficiency than ever before.

Within the context of these evolving social realities, diverse investigators have examined the influence of individual, social and instructional factors on second-language learning. They have attempted to define more precisely "what it is" that the student acquires and how this changes over time as a result of continued instruction or exposure to a community of target-language speakers.

In this paper, I propose to look retrospectively within the broad framework of language and educational planning at the second-language research and theorizing of the past decade; to comment on its contribution; to allude to its application; and to suggest existing lacunae.

During the past decade there has occurred both at home and abroad a great deal of research, theorizing and discussion about various facets of language learning and language teaching. The continuing dialogue has involved at various times parents, educators, researchers and even policy makers at the highest levels. This really should not seem surprising to you if you stop to consider that there are many more bilinguals in the world than monolinguals and that there are many more students who by choice or by necessity attend schools where the medium of instruction is their second or later-acquired language than those who attend schools taught in their mother tongues.

In the United States, much of the work has been prompted by the challenge to provide effective second language teaching within the context of a public school system while simultaneously nurturing the native language development and sociocultural traditions of increasingly large and heterogeneous student populations. According to Waggoner (1976) approximately 25.3 million individuals aged 4 or older live in American households in which languages other than English are spoken. Of this group, approximately 74 per cent (i.e., more than 18 million) speak languages other than English themselves as their usual or second language—a fact which simply astonishes many of my overseas colleagues. The challenge, in this case, appears to be leading to the implementation of bilingual education programs for non-English speakers in many communities (see, for example, Alatis & Twaddell, 1976; Troike & Modiano, 1975). It is interesting to observe, however, that as yet proportionately few English mother tongue speakers have participated in such programs and furthermore that there exists widespread concern in the U.S. about the general apathy toward foreign language study.

In Canada, the necessity exists to develop a citizenry of whatever ethnic origin who are functionally bilingual in English and in French. The policy of the Federal government through the Official Languages Act explicitly advocates the encouragement of multiculturalism through a policy of bilingualism. Unless bilingualism or at least appreciation for bilingualism can be made a fact of life throughout the country (and perhaps, pessimistically even if it can), the present political structure will not survive. Although drastic steps have been taken within the last five years to upgrade the second language skills of civil servants at all levels, the effective development of a bilingual citizenry will probably necessitate major educational reform such as that now occurring within English school boards in many areas of Quebec and Ontario and the passing of several generations.

In many other countries of the world—particularly those of the so-called Third World—a very different situation seems to exist. With the advent of universal primary education, the desire to ensure literacy in the mother tongue, to maximize the relevance of education for those who will likely not complete their schooling and to enhance national pride and unity probably far overshadow the goal of developing the highest possible level of second language proficiency in a select elite. Thus, in many countries, decisions have recently been taken to *decrease* the use of a foreign language as a medium of instruction in that country's schools (e.g., Algeria, the Philippines, the Sudan) or even to *decrease* the amount of time devoted to teaching the foreign language as a subject (e.g., Jordan, the Sudan). Despite such decisions, many individuals have a demonstrated need for greater foreign language proficiency than ever before (see, for example, Harrison, Prator & Tucker, 1975; Tucker, 1977). English or other second-language teachers in these countries face a grave challenge.

It is within the context of evolving social realities such as these that diverse investigators have examined the influence of individual, instructional and social factors on second language learning. They have attempted to define more precisely "what it is" that the student or learner acquires and *how* as well as *why* this changes over time as a result of continued instruction or exposure to a community of target-language speakers.

I have organized the discussion to follow around three general themes: 1) national policy, educational planning and language teaching; 2) the specification of language-teaching objectives within the constraints of national or local policy; and 3) the implementation of language teaching policies. Whenever appropriate, I have discussed the role of research, evaluation or measurement which comprise important elements of each of these themes.

National Policy, Educational Planning and Language Teaching

I want to begin by suggesting that it is necessary to examine various aspects of second language learning and teaching within the broad framework of educational planning. The selection of a language to be taught or to be used as a medium of instruction clearly constitutes an important aspect of educational and of national planning (see, for example, Fishman, 1974). I have come to believe over the last ten years on the basis of extensive experiences in North America, the Middle East and southeast Asia that despite research, experimentation or innovation second language teaching programs or bilingual education programs will *not* succeed or thrive unless they are consistent with government policy, whether explicit or implicit, or with the carefully and clearly expressed goals of local educational authorities (see, for example, Harrison, Prator & Tucker, 1975; Tucker, 1977). Educational or national policy serves to define the parameters within which language teaching programs can be developed.

Let me present briefly a few examples. Consider first the Sudan—a linguistically diverse, geographically immense country with the potential, as yet unrealized, to become a major agricultural resource for the Arab World. During the time of the Anglo-Egyptian Sudan, English was the medium of instruction in many of the schools throughout the country. More recently, a number of small but incrementally rather dramatic changes have occurred which have eroded the official position of English within the country but not diminished its importance. In the northern provinces of the country, Arabic is now the exclusive medium of instruction. The introduction of English as a subject for study was cut back from grade 5 to grade 7; the number of periods per week devoted to English instruction was reduced by two; the length of individual periods for all subjects was reduced; English was declared to be no longer a required "pass" subject for the secondary school leaving examination; *but* English remains the medium of instruction at the University of Khartoum—a fine institution with a mixture of Sudanese and expatriate professors and a continuing British tradition of utilizing the services of external examiners for degree candidates. It has long been rumored that the university will eventually arabicize but that probably will not happen for some years yet.

The situation in the north of the Sudan contrasts sharply with that in the three southern provinces. In the south, English continues to be the major link language and, in fact, is used in many different situations to link the north and the south. In the south, it has been decided to use one of (at least nine) the local vernaculars as the initial medium of primary education with English and Arabic to be taught as second languages and English to be used as the medium of instruction at the secondary level. This radical difference in policy between north and south represents a major concession by the north—a concession articulated in the Addis Ababa agreement of 1972 which concluded a seventeen-year civil war in which language was one of the precipitating factors. If the country is to survive politically linguistic diversity must not only be tolerated, but encouraged.

In the Sudan, then, a series of major educational and national policy decisions have affected the course of English language teaching leading on the one hand to desperation on the part of some northern teachers who must now train students under circumstances relatively less hospitable than those a decade ago and on the other hand to optimism on the part of southern teachers who see once again official sanction and concern for their efforts.

Let's switch for a moment to Canada—in particular to the Province of Quebec. With the passage in 1974 of the Official Language Act, French became the sole official language of the province (but not, it must be remembered, of the country). A number of associated regulations were promulgated to insure that French became the *de facto* as well as the *de jure* language. For example, demonstrable proficiency in French is now a requisite for membership in professional groups such as the corporation of dentists or psychologists or physicians. In addition, companies are required to obtain "francization" certificates to be eligible to compete for government contracts, subsidies, etc. To qualify, French must not only be the actual working language of the organization, but francophones must also be represented at various levels of the company hierarchy including upper management positions.

Perhaps, the single most important and controversial provision of the Act, however, was that which limited access to English-medium schools to anglophones or to other pupils who *already* possessed a "sufficient knowledge" of English to profit from instruction in the language. The government's decision to implement this section of the Act by utilizing examinations of dubious validity and reliability for screening purposes has resulted in the launching of legal action by groups representing both immigrant parents and more interestingly French Canadian parents. The passage and implementation of the Act has meant that a majority of the Quebec population has been abruptly denied access to two popular and successful sources of English language training: 1) following some portion of their education in English-medium schools, or 2) working at jobs where English was the *de facto* language of wider communication. Note, in particular, that even French-English bilingual education is specifically prohibited for non-anglophones or non-English dominant immigrants. Despite this government action, the need for English by Quebec residents—anglophone, francophone or other immigrants—who wish to pursue business or trading opportunities with representatives from an increasingly large array of industrialized or developing countries has *not* diminished. The onus of responsibility for meeting this need thus falls on the ESL program within the French educational sector. There is no reason to believe that the program as presently designed and staffed can meet the challenge (see, for example, d'Anglejan, Acheson, de Bagheera & Tucker, 1977).

Returning to the American setting, it is my belief that neither parents nor educators nor policy makers have yet articulated precisely their expectations for the academic, affective, cognitive, linguistic or social development of children who participate in formal education. Quite clearly, there now exists widespread opposition to the simplistic notion that "the purpose of such [bilingual] instruction must be to create English-speaking Americans with the least possible delay" (New York Times, 1976). But what are the policies or goals to be? And who will determine or shape them? In this respect, the TESOL organization as well as the Center for Applied Linguistics have tried to raise the level of individual and governmental awareness concerning the teaching of English as one necessary, but not by itself sufficient, component of a child's total educational program. In fact, the director of CAL, Rudolph Troike, has called for the establishment within the White House of a high level position of Advisor on Language Activities.

The point that I wanted to make through the previous examples is that educational or national policy does describe the strictures within which language teaching programs can be developed. Furthermore, it is obvious that social pressures motivated by a diverse array of contributory factors can lead to policy change. The results *per se* of empirical research—even when widely publicized—however rarely do.

Specification of Language-Teaching Objectives

On the basis of my experience in various settings during the last ten years, I have come to conclude that much more serious attention needs to be given to the task of defining, publicizing and implementing locally appropriate language teaching goals. For example, is it realistic to expect all pupils to develop native-speaker control in each of the four skills of the target language? Should the objective, rather, be for the individual to develop native-like receptive skills together with the communicative ability necessary to express his ideas? How do prospective teachers, employers or other target language speakers react to spoken or written messages that, although comprehensible are marked by various lexical, phonological or syntactic deviations (see, for example, research by Johannson, 1973; Quirk & Svartvik, 1966; Schachter, Tyson & Diffley, 1976)? It may, for example, be perfectly acceptable for an English Canadian member of the Ontario Provincial Parliament who lives in Toronto to comprehend spoken and written French with ease while producing utterances which are markedly deviant. It might, however, be totally unacceptable for an English Canadian member of the Quebec National Assembly to perform similarly. The task of defining a set of locally appropriate goals undoubtedly necessitates an understanding of the sociolinguistic context in which the graduates of the language program will live and work (see Cohen, 1975; Harrison, Prator & Tucker, 1975).

Within the past decade, sociolinguistic or language policy surveys have been conducted in a number of countries. They have been designed to provide information about the patterns of language use within a particular region or country, the aims for language teaching, the dimensions of the language teaching effort, the various resources available to implement the programs—all within the context of the actual demonstrated needs which people have for the target language(s) in diverse daily activities (see, for example, Bender, Bowen, Cooper & Ferguson, 1976; Harrison, Prator & Tucker, 1975; Ladefoged, Criper & Glick, 1972; Royal Commission on Bilingualism and Biculturalism, 1963; Whiteley, 1974). Such surveys are based on the assumption that people's needs for a particular target language should be allowed to influence the scope and design of the language teaching program—an assumption which has certainly not characterized the development of programs in many countries. I have suggested elsewhere (Tucker & d'Anglejan, 1971) that the first step in establishing an innovative language-teaching program "should involve a small-scale sociolinguistic survey of the local community where the program will be situated" (p. 492). In practice, this is rarely done.

But how are language-teaching objectives to be developed? Consider the directive in the "Official Language Act" of the Province of Quebec:

> The curricula must ensure that pupils receiving their instruction in English acquire a knowledge of spoken and written French, and the Ministry of Education shall adopt the necessary measures to that effect.
> The Ministry of Education must also take the necessary measures to ensure instruction in English as a second language to pupils whose language of instruction is French.
> (Bill No. 22, Title III, Chapter V, Article 44)

I realize you may argue that the task of the legislature is simply to provide the mandate under which appropriate experts will articulate more explicitly the goals or objectives of the language teaching program. But consider the set of specific objectives for Jordanian students who by the end of the Compulsory Cycle (i.e., after studying English as a foreign language for five years from grades 5-9), are expected to be able to:

1. Understand simple English spoken at a normal speed.
2. Communicate sensibly with an English-speaking person, within certain reasonable areas.
3. Read simple English with ease, fluency and understanding.
4. Write a basic paragraph in English, using the basic structure of the language (*English Curriculum: Compulsory Stage,* 1969, p. 5).

I have previously claimed (see Harrison, Prator & Tucker, 1975) that "these 'specific' aims typify the conventional goals so often cited for language study. They lack precision, and fail to make clear precisely what degree of English proficiency students are expected to achieve" (p. 2). I went on to suggest that operationally-defined behavioral objectives appeared to be a necessary prerequisite to effective program development and that ultimately the success of an English language program must be measured by the ability of school graduates to obtain suitable employment which would permit them to grow individually and to contribute directly to the social and economic development of their country. In retrospect, my suggestion seems naive and likewise without precision and for the present I shall simply identify this as an area which requires additional thought and development.

Let me conclude this section by drawing your attention to another major factor which shapes the actual language teaching program—the examination system. In many countries, regions, states or provinces the final English language attainment of each school graduate is evaluated by his success or failure on the English paper of a secondary certificate examination (this is true, for example, in Egypt, Jordan, Quebec). It has been my repeated observation that the form and content of this examination affects the manner and substance of English language teaching; and furthermore that even where there exists relatively explicit objectives for ELT there is often a lack of correspondence between the objectives and the examination (see Bending, 1976). In such cases, despite their professional preparation, despite the resources at their disposal, teachers often elect to concentrate their attention on "preparing" their students for the required examination. It is not inconceivable that students who pass the English paper of the Egyptian Thannawiyya amma or of the Quebec matriculation exam can also communicate accurately and confidently at the hypothesis remains for the most part untested. Educators are now beginning to turn their attention toward this domain and I predict that it will become an increasingly important research area.

The Implementation of Language Teaching Policies

Let me now examine somewhat more explicitly factors which may affect the successful implementation of language teaching policies. In recent years, the topic of second language learning and teaching has attracted the attention of an increasing number of researchers drawn from a wide range of academic disciplines. In addition to educators concerned with discovering and applying new and better pedagogical techniques, we find an active group of anthropologists, linguists, psychologists and sociologists engaged in the systematic study of the complex interplay among affective, cognitive, social and other factors in second language acquisition.

A number of scholars have proposed useful models or frameworks within which we can examine questions related to second language learning and teaching (see, for example, Fox, 1975; Paulston, 1974; Schumann, 1976a; Strevens, 1976). I believe that we can represent the domain graphically by using three concentric circles. The outermost represents the *sociocultural* context—the ambiance in which the language is spoken natively, in which it is to be learned as a second or foreign language and is to be used for diverse purposes. It comprises elements such as: 1) the official position and allocated role(s) of the target language in the community, nation or region; 2) the perceived status of the language and its speakers; 3) the existence of structures which encourage or facilitate the use of the target language; 4) the size and cohesiveness of the groups in contact; and 5) the correspondence between their values, attitudes and traditions (c.f., Shumann, 1976b).

The middle circle represents the *instructional* or pedagogical setting—the context in which the language is to be "formally" transmitted from a teacher or other resource person to the learner. Salient components at this level are factors such as: 1) the goals whether explicit or implicit for second language teaching; 2) the pedagogical techniques to be employed; (I am going to restrict my remarks for the most part to second language learning which goes on within the confines of a classroom although I shall allude to issues posed by informal learning; 3) the design of the syllabus including the choice or development and sequencing of materials and the allocation of time; 4) the training and language proficiency of the teacher as well as the attitudes of teachers and administrators; and 5) the procedures for evaluation. The interface between these two circles is effected by educational planning within the context of national or regional priorities.

The innermost circle represents *individual* factors. The second language learner brings with him to the learning environment a variety of attitudes, biological predispositions, learned behaviors and societal experiences that affect the course, the duration and the speed of second language learning. Important factors at this level may be elements such as: 1) age; 2) intelligence and language aptitude; 3) learning style or the general cognitive processes drawn upon; 4) personality characteristics; and 5) motivation and attitude. The interface between the innermost and the middle circle is effected when the learner by choice or by necessity begins to follow one of a large variety of instructional tracks.

The components or elements that I have identified at each level are, of course, only illustrative; they are by no means exclusive. I would like now to raise the general question of how research—specifically that conducted during the last decade or so—has helped us to understand better the process of second language learning and teaching. I will comment briefly on some of the important research trends, research applications and existing lacunae where additional investigations may be warranted. The term "research" should be interpreted very broadly to include, for example, attempts: 1) to describe what it is that the learner acquires during his course of instruction; 2) to posit and verify models to describe the process of second language learning; and 3) to examine the contributions or relationships among the diverse elements described earlier.

I think it would not be inappropriate to begin by observing once again (cf., Tucker & d'Anglejan, 1975, p. 66) that:

> Concurrent with the decline in popularity of the traditional second language methodologies came a radical shift in theoretical views of the language acquisition process.

> Many researchers now view native language learning and second language learning as analogous processes characterized by the development of rule-governed creativity. The learner is thought to engage actively, albeit unconsciously, in the gradual discovery of the rule system underlying the language to which he is exposed. Although the young child acquiring his native language and the second language learner may go about the task in somewhat different ways, the developmental pattern of each is now thought to be both systematic and predictable.

That statement was prompted by the burst of research activity begun during the late 1960's and continuing to the present day which sought, among other things, to describe various facets of the speech system acquired by second language learners. Let me for purposes of this discussion over-simplify grossly the scenario of this period. The efficacy of language teaching programs based upon notions derived from behaviorist views of learning which stressed the role of the mimicry, practice and memorization of discrete units of language chosen on the basis of *a priori* contrastive analyses and presented in a structurally-graded sequence was seriously questioned. Disenchantment with student achievement was rampant.

As a result of provocative writing and research by Noam Chomsky, his students, his colleagues and their students, the operational definition of language learning underlying most recent research shifted from one which considered the child's linguistic growth to be the result of reinforcement or shaping by his environment to what Brown (1973) has described as a creative construction process. Child language was no longer viewed as a defective or haphazard form of adult speech but rather as an orderly, rule-governed system which evolves through a series of predictable stages toward eventual adult competence.

The incisive notion that second language learning, even within the classroom context, might reveal patterns and regularities similar to those found in native language learning had, of course, been posited by scholars such as Corder, (1967), Nemser, (1971), Selinker, (1972) and more recently by Selinker, Swain and Dumas (1975). Their speculations were supported by a spate of "error analysis" studies (see, for example, Richards, 1973; 1974; Scott & Tucker, 1974; Taylor, 1975; but see also Hammarbarg, 1974). Researchers observed that in addition to errors attributable to interference from the learner's mother tongue (which seem especially prominent at the elementary stages of instruction) there are other common systematic errors which can *not* be attributed to interference and which are found in the speech of learners from a variety of language backgrounds studying the same target language. A plausible inference is that these latter error types result from the application of general operating principles, learning strategies or cognitive processes to the task at hand (cf., Hatch, 1974; Reibel, 1971; Sampson & Richards, 1973; Tarone, Cohen & Dumas, 1976). An important pedagogical notion which derives from this work is that second language teachers should come to regard "errors" as indicators of progress while at the same time continuing to provide the student with well formed and varied models. I believe it is important in this regard *not* to confuse the notion that errors may be used for diagnostic purposes with a general license for permissiveness that would imply a lack of concern with the goal of achieving facility in the standard form of the target language as well as in other socially, regionally or situationally appropriate varieties.

One important direction to this research has been the attempt to define more precisely *what* it is that the student (or the "free" learner) acquires and *how* as well as *why* this changes over time as a result of continued instruction or exposure to a community of target language speakers. Thus, a broad ranging series of studies have been conducted to examine the acquisition of various features of the second language such as grammatical functors or morphemes (e.g., Bailey, Madden & Krashen, 1974; Dulay & Burt, 1974; Hakuta, 1974; Larsen-Freeman, 1975), the auxiliary system (Cancino, Rosansky & Schumann, 1975), negation (Cazden, Cancino, Rosansky & Schumann, 1975; Milon, 1974) *wh* questions (Ravem, 1974); and other perhaps "higher-order" syntactic structures (e.g., d'Anglejan & Tucker, 1975; Boyd, 1975; Cohen, 1976; Cook, 1973; Gillis & Weber, 1976; Hamayan, Markman, Pelletier & Tucker, 1976; Krashen, Sferlazza, Feldman & Fathman, 1976) as well as phonology (Tarone, 1974; 1976). Researchers have attempted to draw inferences from a variety of cross-sectional or longitudinal studies conducted with learners from different backgrounds about the effects of factors such as age, formality and type of language training and frequency of exposure, on second language acquisition (cf., Fathman, 1975; Larsen-Freeman, 1976a; Wagner-Gough & Hatch, 1975).

The results of the research completed to date indicate that the performance of the learner is not random or disorganized but there are, at this time, few other compelling pedagogical implications. For example, the order of acquisition of grammatical morphemes may or may not reflect accurately general second language development, may or may not be similar for first or second language learners, may or may not be similar for learners acquiring their second language formally or informally and may or may not permit us to draw valid inferences about the operating principles or strategies which the learner applies.

I would argue strongly (as do Tarone, Swain & Fathman, 1976) that it is premature to consider revising syllabi or instructional approaches to try to reflect the results of these studies (but see Taylor, 1975b); although I will shortly argue that there are certain broad areas of application for the results of second language research.

Let me summarize this present section by observing that I view this area of research—that of attempting to define more precisely what it is that the learner acquires and how and why this product changes over time—to be of central importance for the development of effective second language teaching programs. A clearer understanding of the discrepancy between well-articulated goals for second language teaching and the typical product should aid our diagnostic procedures and facilitate the introduction of corrective action when and where appropriate.

Let me return now to a brief consideration of some other recent representative research which has examined the contributions or relationships among factors or elements of the three concentric circles which I earlier used to define the context for second language learning and teaching.

At the outermost or *sociocultural* level, I have already described in my introductory section the importance of language planning activities or sociolinguistic surveys as tools to provide baseline information necessary for the rational development of language teaching programs. I will not belabor this discussion except to note the important contributions in this area by scholars such as Joshua Fishman and his colleagues (Fishman *et al.*, 1966; Fishman, Cooper & Ma, 1971), Wallace Lambert (1973) and Stanley Lieberson (1970) who have provided continuing reminders that language learning occurs within a social context, that our research strategies must involve attempts to examine the values, attitudes, traditions, political factors and socioculural dynamics that underlie language learning and that our research must be interdisciplinary.

Let me turn now to the innermost or *individual* level. A plethora of studies have been conducted which seek to examine, to measure or to manipulate a diverse array of individual characteristics. For example, researchers have investigated in various permutations the relationship among age (cf., Krashen, 1973; Ramsey & Clark, 1974), affective factors (cf., H. D. Brown, 1973; Chastain, 1975; Gardner & Lambert, 1972; Gardner, Smythe, Clément & Gliksman, 1976; Jacobson & Imhoof, 1974; Taylor, 1974; Teitlebaum, Edwards, Hudson & Hudson, 1975; Schumann, 1975), "ability" factors (cf., Carroll, 1974; Lambert & Gardner, 1972) learning stages or style (cf., Naiman, Fröelich & Stern, 1975; Rosansky, 1974) and second language learning. Two general conclusions can be drawn from these studies: 1) aptitude and attitude/motivation comprise statistically independent and significant predictors of language learning success; and 2) youthful vs. adult learners *may* approach the language learning task differently and hence profit from different teaching approaches. The influences of personality variables and differences in cognitive or learning style are as yet little understood and research should be continued in these areas. The major application of these studies derives from the fact that teachers can affect or manipulate attitude and motivation and can design programs of instruction appropriate to the age, experience, ability and interest of their students. The apparent popularity and success of "English for Special Purposes" programs lends credence to this claim (see, for example, ESPMENA).

During the past two years, a few researchers working at this level have begun to phrase their questions and shape their data gathering techniques slightly differently. They are attempting to define the constellation of individual factors associated with successful second language learning and furthermore to determine whether the effects of these factors can be modulated by the language teaching program which a student is following (cf., Cohen, in press; Naiman, Fröelich & Stern, 1975; Rubin, 1975; Tucker, Hamayan & Genesee, 1976). The results of the few studies conducted to date indicate that the study of individual differences in second language learning and factors associated with them can be usefully pursued using multivariate statistical procedures. This area, I believe, constitutes an important future research direction.

Let me turn finally to the remaining level—the *instructional* setting and its associated factors. This level has, I think, been profoundly affected by the theoretical insights and empirical research results of the past decade. One of the important pedagogical notions which derives from insights about native language learning is that the major focus of second language classroom activity should be on communication—not just of simulated dialogues—but no genuine communication where the validity of students' utterances will be judged on the basis of their content rather than on the appropriateness of their grammatical form (Diller, 1975; Tucker & d'Anglejan, 1975). The apparent acceptance of this notion manifests itself in diverse ways: 1) by the adoption of seemingly radical objectives and techniques on the part of some teachers (e.g., the development of communicative competence at the possible expense of grammatical perfection, Savigon, 1972); 2) by the switch from structurally-graded to notionally-based syllabi (O'Neill, 1976); 3) by a widespread move toward the implementation of apparently successful bilingual educational programs in many parts of North America (see Alatis & Twaddell, 1976; Troike & Modiano, 1975); and 4) by the development of English-for-special-purposes texts and courses throughout much of the third world where English occupies the role of an important foreign or second language. Such innovations all attempt to capitalize on the observation that a student can effectively acquire a second language when the task of learning the language becomes incidental to the task of communicating with someone about something which is inherently interesting.

This shift in focus calls, of course, for a reassessment of the role and training of the second language teacher—a reassessment which at least one major professional organization has recently conducted (TESOL, 1976). The second language teacher will often be called upon to work in close partnership with teachers of content subjects (see, for example, Cohen, 1975b). The language specialist must not only be a fluent speaker of the target language; but should also during the course of professional preparation have acquired an understanding of current theories of language acquisition and be familiar with and attuned to the sociocultural traditions of the students. It should be emphasized that there remain enormous variations and healthy controversy about the optimal manner of implementing programs of formal instruction for children or for adults (cf., Diller, 1975; Krashen & Seliger, 1975) today as there were fifty years ago (Jesperson, 1904; Palmer, 1921).

Research continues unabated at this instructional level, but I believe we are only now beginning to ask relevant questions. We have only recently begun to examine minutely what actually goes on in the language classroom (Shultz, 1975); what input the student receives and how this influences second language acquisition (Hatch, Shapira & Gough, 1975; Larsen-Freeman, 1976b; Keller-Cohen & Gracey, 1976; Wagner-Gough & Hatch, 1975). So there remains much to be accomplished.

Concluding Remarks

It remains simply to make a series of summary statements to reiterate what I see as the major themes running through the research that I have reviewed.

1. Language teaching programs occur with a sociocultural context. They are shaped by and contribute to national and educational planning;
2. Serious attention needs to be given to the task of defining, publicizing and implementing locally appropriate language teaching goals;
3. In many areas a serious lack of correspondence between the objectives of English language teaching programs and the form and content of the national examination system results in ineffective or misguided teaching practices;
4. An understanding of second language learning and teaching necessitates a consideration of the interaction among sociocultural, instructional and individual factors;
5. The second language learner like the child native speaker is thought to engage actively, albeit unconsciously, in the discovery of the rule system underlying the language to which he is exposed;
6. The developmental pattern of the learner appears to be systematic and predictable;
7. We are beginning to describe and to understand what it is that a student acquires when he learns a language as well as how and why this is modified over time; but our descriptive tools and analytical insights remain relatively naive;
8. It appears to be premature to consider revising syllabi to try to reflect the results from these initial descriptive studies;
9. The notion that the major focus of activity in the classroom should be on meaningful communication appears to be theoretically and empirically well founded;
10. Teachers can affect the attitudes and motivation of their students and can contribute to the design of programs of instruction appropriate to the age, ability, experience and interest of their students;
11. An examination of the constellation of factors associated with language learning using multivariate techniques is a useful research direction to pursue as is a closer examination of the dynamics of classroom language usage and patterns of interaction;
12. The theoretical validity of a particular approach and considerations of methodological sophistication may be relatively minor ingredients in the

development of a successful and total second language education program when compared with sociocultural or attitudinal considerations.

In summary, I know of no "scientific" data or empirical evidence which demonstrate conclusively that we have yet designed the ideal program or programs by which a second language can be *taught*; but we do know, that almost any student (and we see ample evidence) placed in the appropriate milieu, given ample opportunity and the support of his parents, peers and teachers can successfully *acquire* a second language—the paradox may be semantic or illusory but clearly the challenge to provide effective second language training to an increasingly large and heterogeneous group of students remains.

REFERENCES

d'Anglejan, A., and G. R. Tucker. 1975. The acquisition of complex English structures by adult learners. *Language Learning 25*, 281–296.

d'Anglejan, A., P. Acheson, I. de Bagheera, and G. R. Tucker. 1977. English as the second language of Quebec: a teacher profile. Paper presented at TESOL convention, Miami Beach.

Alatis, J. E. and K. Twaddell, (eds.) 1976. *English as a second language in bilingual education*. Washington, TESOL.

Bailey, N., C. Madden, and S. D. Krashen. 1974. Is there a "natural sequence" in adult second language learning? *Language Learning 24*, 235–243.

Bender, M. L., J. D. Bowen, R. L. Cooper, and C. A. Ferguson, (eds.) 1976. *Language in Ethiopia*. London, Oxford University Press.

Bending, H. B. 1976. Motivation in an examination—geared school. *English Language Teaching 30*, 315–320.

Boyd, P. A. 1975. The development of grammar categories in Spanish by Anglo children learning a second language. *TESOL Quarterly* 9, 125–135.

Brown, H. D. 1973. Affective variables in second language acquisition. *Language Learning 23*, 231–244.

Brown, R. 1973. *A first language: the early stages*. Cambridge, Mass., Harvard University Press.

Cancino, H., E. J. Rosansky, and J. H. Schumann. 1975. The acquisition of the English auxiliary by native Spanish speakers. *TESOL Quarterly* 9, 421–430.

Carroll, J. B. 1975. Aptitude in second language learning. *Proceedings of the Fifth Symposium of the Canadian Association of Applied Linguistics*.

Cazden, C. B., H. Cancino, E. J. Rosansky, and J. H. Schumann. 1975. Second language acquisition sequences in children, adolescents and adults. Final Report Project No. 730744 grant No. NE-6-00-3-0014. U.S. Dept. of H, E & W, NIE.

Chastain, K. 1975. Affective and ability factors in second-language acquisition. *Language Learning 25*, 153–161.

Cohen, A. D. 1975a. *A sociolinguistic approach to bilingual education*. Rowley, Mass., Newbury House.

Cohen, A. D. 1975b. Error correction and the training of language teachers. *Modern Language Journal 59*, 414–422.

Cohen, A. D. 1976. The acquisition of Spanish grammar through immersion: some findings after four years. *The Canadian Modern Language Review* 32, 562–574.

Cohen, A. D. In press. Successful second language speakers: a review of the research literature. *Israeli Journal of Applied Linguistics*.

Cook, V. J. 1973. The comparison of language acquisition in native children and foreign adults. *IRAL 11*, 13–28.

Corder, S. P. 1967. The significance of learner's errors. *IRAL 5*, 161–170.

Diller, K. C. 1975. Some new trends for applied linguistics and foreign language teaching in the United States. *TESOL Quarterly 9*, 65–73.

Dulay, H., and M. Burt. 1974. A new perspective on the creative construction processes in child second language acquisition. *Language Learning 24*, 253–278.

ESPMENA, English for special purposes in the Middle East and North Africa, English Language Servicing Unit, University of Khartoum, Sudan.

Fathman, A. 1975. The relationship between age and second language productive ability. *Language Learning 25*, 245–253.

Fishman, J. A. 1974. Language modernization and planning in comparison with other types of national modernization and planning. In J. A. Fishman (ed.) *Advances in language planning* The Hague, Mouton. Pp. 79–102.

Fishman, J. A. *et al.* 1966. *Language loyalty in the United States*. The Hague, Mouton.

Fishman, J. A., R. L. Cooper, and R. Ma, *et al.* 1971. *Bilingualism in the barrio*. Bloomington, Indiana University Press.

Fox, M. J. 1975. *Language in education: problems and prospects in research and training*. New York, Ford Foundation.

Gardner, R. C., and W. E. Lambert. 1972. *Attitudes and motivation in second language learning*. Rowley, Mass., Newbury House.

Gardner, R. C., P. C. Smythe, R. Clément, and L. Gliksman. 1976. Second language learning: a social psychological perspective. *Canadian Modern Language Review 32*, 198–213.

Gillis, M., and R. Weber. 1976. The emergence of sentence modalities in the English of Japanese-speaking children. *Language Learning 26*, 77–94.

Hakuta, K. 1974. A preliminary report on the development of grammatical morphemes in a Japanese girl learning English as a second language. *Working Papers in Bilingualism 3*, 18–43.

Hamayan, E., B. R. Markman, S. Pelletier, and G. R. Tucker. 1976. Differences in performance in elicited imitation between French monolingual and English-speaking bilingual children. *Working Papers in Bilingualism 8*, 30–58.

Hammarbarg, B. 1974. The insufficiency of error analysis. *IRAL 12*, 185–192.

Harrison, W., C. Prator, and G. R. Tucker. 1975. *English-Language Policy Survey of Jordan*. Arlington, Va., Center for Applied Linguistics.

Hatch, E. 1974. Second language learning universals? *Working Papers in Bilingualism 3*, 1–17.

Hatch, E., R. Shapira, and J. Gough. 1975. "Foreigner-talk" Discourse. UCLA, August 1975. Mimeo.

Jacobsen, M., and M. Imhoof. 1974. Predicting success in learning a second language. *Modern Language Journal 55*, 329–336.

Jespersen, O. 1904. *How to teach a foreign language*. London, George Allen & Unwin Ltd.

Johansson, S. 1973. The identification and evaluation of errors in foreign languages: a functional approach. In J. Svartvik (ed.) *Errata: papers in ERROR analysis*. Lund, Sweden, CWK Gleerup. Pp. 102-114.

Jordan, Ministry of Education, Curricula and Textbooks Division. 1969. *English Curriculum: Compulsory Stage*. Amman, Cooperative Printing Presses Workers Society, 2nd ed.

Keller-Cohen, D., and C. Gracey, 1976. Repetition and turn allocation in the non-native acquisition of discourse. Paper presented at Eighth Child Language Research Forum, Stanford, April 1976.

Krashen, S. 1973. Lateralization, language learning, and the critical period: Some new evidence. *Language Learning 23*, 63–74.

Krashen, S. D., and H. W. Seliger, 1975. The essential contributions of formal instruction in adult second language learning. *TESOL Quarterly 9*, 173–183.

Krashen, S. D., V. Sferlazza, L. Feldman, and A. K. Fathman. 1976. Adult performance on the SLOPE test: more evidence for a natural sequence in adult second language acquisition. *Language Learning 26*, 145–151.

Ladefoged, P., R. Glick, and C. Griper. 1972. *Language in Uganda*. Nairobi, Oxford University Press.

Lambert, W. E. 1973. Culture and language as factors in learning and education. Paper presented at the fifth annual learning symposium at Western Washington State College; Bellingham, Washington; November, 1973.

Larsen-Freeman, D. E. 1976a. An explanation for the morpheme acquisition order of second language learners. *Language Learning 26*, 125–134.

Larsen-Freeman, D. 1976b. ESL teacher speech as input to the ESL learner. *UCLA Workpapers in Teaching English as a Second Language 10*, 45–50.

Larsen-Freeman, D. E. 1975. The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly* 9, 409–419.

Lieberson, S. 1970. *Language and ethnic relations in Canada*. New York, Wiley.

Milon, J. P. 1974. The development of negation in English by a second language learner. *TESOL Quarterly* 8, 137-144.

Naiman, N., M. Fröhlich, and H. H. Stern. 1975. The good language learner. Modern Language Centre Ontario Institute for Studies in Education, Multilith.

Nemser, W. 1971. Approximative systems of foreign language learners. *International Review of Applied Linguistics*, 9, 115–123.

New York Times. 1976. Bilingual Danger, Editorial, November 22, 1976.

O'Neill, T. 1976. The teaching of notions. *In English language teaching for the Arab world*. London, Oxford University Press. Pp. 10–15.

Palmer, H. E. 1921. *The oral method of teaching languages*. Cambridge, W. Heffer & Sons Ltd.

Paulston, C. B. 1974. Implications of language learning theory for language planning: concerns in bilingual education. Papers in Applied Linguistics, Bilingual Education Series: 1. Arlington, Va.: Center for Applied Linguistics.

Quirk, R., and J. Svartvik. 1966. *Investigating linguistic acceptability*. The Hague, Mouton.

Ramsey, C. A., and E. N. Wright. 1974. Age and second language learning. *The Journal of Social Psychology 94*, 115–121.

Ravem, R. 1974. The development of wh-questions in first and second language learners. In J. H. Schumann & N. Stenson (eds.) *New frontiers in second language learning*. Rowley, Mass., Newbury House.

Reibel, D. A. 1971. Language learning strategies for the adult. In P. Pimsleur & T. Quinn (eds.) *The psychology of second language learning*. Cambridge: University Press, Pp. 87-96.

Richards, J. C. 1973. A noncontrastive approach to error analysis. In J. W. Oller & J. C. Richards (eds.) *Focus on the learner: pragmatic perspectives for the language teacher*. Rowley, Mass., Newbury House.

Richards, J. C. 1974. *Error analysis: perspectives on second language learning*. London, Longmans.

Rosansky, E. 1975. The critical period for the development of language: some cognitive developmental considerations. *Working Papers on Bilingualism 6*, 92–102.

Royal Commission on Bilingualism and Biculturalism. 1963. Ottawa, Queen's Printer.

Rubin, J. 1975. What the "Good language learner" can teach us. *TESOL Quarterly* 9, 41–51.

Sampson, G. P., and J. C. Richards. 1973. Learner language systems. *Language sciences*, August, 18–25.

Savignon, S. 1972. *Communicative competence: an experiment in foreign-language teaching*. Philadelphia, Center for Curriculum Development.

Schachter, J., A. F. Tyson, and F. J. Diffley. 1976. Learner intuitions about grammaticality. *Language Learning 26*, 67–76.

Schumann, J. H. 1976a. Second language acquisition research: getting a more global look at the learner. *Language Learning* Special Issue no. 4 (Papers in Second Language Acquisition), 15–28.

Schumann, J. H. 1976b. Social distance as a factor in second language acquisition. *Language Learning 26*, 135–143.

Schumann, J. H. 1975. Affective factors and the problem of age in second language acquisition. *Language learning* 25, 209-235.

Scott, M. S., and G. R. Tucker. 1974. Error analysis and English-language startegies of Arab students. *Language Learning 24*, 69-97.

Selinker, L. 1972. Interlanguage. *IRAL 10*, 209–231.

Selinker, L., M. Swain, and G. Dumas. 1975. The interlanguage hypothesis extended to children. *Language Learning* 25, 139–152.

Shultz, J. 1975. Language use in bilingual classrooms. Paper presented at TESOL Convention, Los Angeles.

Strevens, P. 1976. A theoretical model of the language learning/language teaching process. *Working Papers in Bilingualism 11*, 129-152.

Tarone, E. 1974. Speech perception in second language acquisition: a suggested model. *Language Learning* 24, 223–233.

Tarone, E. 1976. Some influences on interlanguage phonology. *Working papers in bilingualism* 8, 87–110.

Tarone, E., A. D. Cohen, and G. Dumas. 1976. A closer look at some interlanguage terminology: a framework for communication strategies. *Working papers on Bilingualism* 9, 76–90.

Tarone, E., M. Swain, and A. Fathman. 1976. Some limitations to the classroom applications of current second language acquisition research. *TESOL Quarterly 10*, 19–32.

Taylor, B. P. 1974. Toward a theory of language acquisition. *Language Learning 24*, 23–35.

Taylor, B. P. 1975a. The use of overgeneralization and transfer learning strategies by elementary and intermediate students of ESL. *Language Learning* 25, 73–107.

Taylor, B. P. 1975b. Adult language learning strategies and their pedagogical implications. *TESOL Quarterly* 9, 391–399.

Teitelbaum, H., A. Edwards, and A. Hudson. 1975. Ethnic attitudes and the acquisition of Spanish as a second language. *Language Learning* 25, 255–266.

TESOL Position paper on the role of English as a second language in bilingual education, March, 1976.

Troike, R. C., and N. Modiano. 1975. *Proceedings of the First Inter-American Conference on Bilingual Education*. Arlington, Va., Center for Applied Linguistics.

Tucker, G. R. 1977. Some observations concerning bilingualism and second-language teaching in developing countries and in North America. In P. A. Hornby (ed.) *Bilingualism: psychological, social and educational implications*. New York, Academic Press. Pp. 141–146.

Tucker, G. R., and A. d'Anglejan. 1971. Some thoughts concerning bilingual education programs. *The Modern Language Journal* 55, 491–493.

Tucker, G. R., and A. d'Anglejan. 1975. New directions in second language teaching. In R. C. Troike and N. Modiano (eds.) *Proceedings of the First Inter-American Conference on Bilingual Education*. Arlington, Va., Center for Applied Linguistics. Pp. 63–72.

Tucker, G. R., E. Hamayan, and F. H. Genesee. 1976. Affective, cognitive and social factors in second language acquisition. *The Canadian Modern Language Review* 32, 214–226.

Waggoner, D. 1976. NCES' Survey of languages. *Linguistic Reporter 19*, 3, 5 & 8.

Wagner-Gough, J., and E. Hatch. 1975. The importance of input data in second language acquisition studies. *Language Learning* 25, 297–308.

Whiteley, W. H. (ed.) 1974. *Language in Kenya*. Nairobi: Oxford University Press.

Bilingual Education and ESL: A Symbiotic Relationship

Hernan LaFontaine

Center for Bilingual Education
New York City Board of Education

Recent events have stirred anew an alleged conflict between programs in the teaching of English as a Second Language and bilingual education programs. Whatever the motivation for the revival of this misconception, ESL and bilingual education are not only compatible but essential to the effective implementation of both approaches. The elimination of this mythical conflict can be best achieved by clarifying the objectives of both programs revealing the mutual benefits that can be derived from a close cooperation between practitioners in both fields. A comprehensive effort to implement bilingual education programs in New York City resulting from a court order offers evidence that ESL instruction is increasing rather than decreasing. Our precious and limited time and effort is more productively expended in improving our practices rather than engaging in unnecessary and wasteful accusations. The National Association for Bilingual Education and TESOL have already undertaken various steps to emphasize the cooperation and support of both organizations for ESL and bilingual education programs.

As a former biology teacher I could not resist Jim Alatis' reference to the concept of symbiosis in his address to this convention in 1975 on "The Compatibility of TESOL and Bilingual Education." To me, it seems to focus on the very essence of our continued efforts to unify our profession in the struggle to develop and support high quality programs which provide equal education opportunity, which nurture academic excellence and which attempt to move our society towards the notion of cultural democracy. In its biological context symbiosis refers to a phenomenon in which two living things function in a relationship which is mutually beneficial. It differs from parasitism in which only one party benefits from the relationship and it certainly differs from a saprophytic relationship in which one party feeds on the other until the other is dead. To bilingual educators, therefore, I am anxious to announce that, in spite of recent opinion to the contrary, ESL is *NOT* hazardous to your health. I feel a little like the Surgeon General of the United States making an announcement about a controversial drug. Can you imagine comments such as: "You're not taking ESL are you?" or "I think those kids are on ESL." But I can easily sympathize because there are those that characterize bilingual education in a similar manner. Without mentioning the name, one state refers to their program as "terminal bilingual education." Poor kids! Inflicted with such a terrible disease at such a young age! "Madam, I'm sorry to inform you but your son has—BILINGUALITIS." I repeat, ESL is not hazardous to your health, but to teachers of ESL I hasten to add, *when administered properly*. Similar to the bilingual education "pill", the dosage, potency, and quality are crucial otherwise the tongue begins to ossify and the result is that the patient or student can't speak either language. The well worn proverb that says that a lot of truth is said in jest seems appropriate here.

I hope today to make my modest contribution towards laying to rest any thought of real or imagined conflict between the ESL and bilingual education approaches and between their practitioners. It seems that everytime we think we're making progress toward this end something happens which starts the quarrel all over again. I am referring, for example, to the U.S. Commission on Civil Rights' report "A Better Chance to Learn: Bilingual Bicultural Education" and to Dr. Troike's, now famous, editorial in the Linguistic Report. In both cases the underlying premise seems to be to consider ESL and bilingual education as mutually exclusive and to present the issues in such a way as to attempt to force practitioners and consumers to choose one or the other. Mary Hines' critique of the Commission's report and Dr. Spolsky's and Dr. Kaplan's response to Dr. Troike's statement both clearly analyze and expose the fallacy of such a premise. Spolsky cautions us not to throw the baby out with the bath water. Hines urges us to do more research on ways to improve what we are doing. I agree wholeheartedly with both and can only add the emphasis that none of us believes that either ESL or bilingual education is a panacea which will cure all of the educational ills which beset our youngsters. On the contrary, we are constantly searching for improved techniques and materials to strengthen our instructional programs. And most often we do this together in a cooperative approach. It is sheer nonsense to believe that ESL and bilingual education are totally unrelated. No one ever asks whether we should teach geometry or mathematics; or biology or science! Such questions would be met with astonished laughter. Perhaps we should respond in the same way. But, keeping in mind that those who do pose the question are often ignorant of our objectives we should take the time to educate them to the intricacies of our efforts.

The ESL program goal of having students acquire a second language is by definition an integral goal of bilingual education. Otherwise, *bi*-lingual education could not be characterized as being bi-lingual! Please pardon such a simplistic and unsophisticated argument before this audience but I can tell you that I have had to regress to such elementary logic many times in an effort to convince detractors and other instigators that we are all trying to accomplish the same goals. Bilingual educators in the United States are all in agreement that it is crucial to academic and career success of our students that they master the English language. We should point out too that in most cases we are also trying to reach the same group of youngsters. In other words, our present priority in bilingual education is to serve the student who is limited in his ability to function in English and, of course, this is the same student who will benefit from instruction in ESL.

The added features of a bilingual program focus on our concern with providing the student with an opportunity to participate in a meaningful way in the rest of the instructional program. Our contention is that this can best be accomplished by using the student's native language as a medium of instruction, by continuing to develop his skills in his native language and by helping him to learn and understand his cultural background as a component of a multicultural society. This last concern has clear implications for the possibility of helping the student develop a positive self-image and hopefully the self-confidence necessary to motivate an aspiration of success in school and in life.

In all of TESOL's publications and in many ESL conferences I have attended I have yet to see or hear any opposition to these basic concepts of bilingual education. We must continue to emphasize this agreement between us as an example of the compatibility between ESL and bilingual education.

And now if you will permit me a brief "commercial" I would like to very briefly describe our situation in New York City as a further example of this compatibility. You may know already that the N.Y.C. Board of Education is implementing a bilingual program as a result of a consent decree signed with ASPIRA of New York and approved by a federal court judge. The Consent Decree program, as we refer to it to distinguish it from other bilingual programs we offer, specifically mandates that "a planned and systematic program designed to develop the child's ability to speak, understand, read and write the English language" be provided to all eligible students. As a result, I am informed by David Krulik head of the ESL unit in our Office of Bilingual Education, that more students are participating in ESL instruction than ever before. In other words, the court mandate for bilingual programs has included in a practical sense a mandate for ESL instruction. This mandate has led to a need for more ESL teachers, so that this year in spite of continued financial restrictions and subsequent dismissal of teachers, no ESL teachers have been laid off. In fact, it appears that ESL teachers will continue to be appointed to serve in the Consent Decree program. Now, don't all of you run to New York looking for employment. To be appointed you must possess a New York City license, as a Teacher of English as a Second Language. And whether we agree or not with the New York City teacher licensing procedure we certainly agree that an ESL teacher must be specially trained to provide such instruction. To say that any teacher who speaks English is automatically equipped to teach ESL is as absurd as saying that any teacher who speaks Spanish can teach any technical subject in Spanish. Lack of time and the extensive literature already available obviate the necessity for further elaboration on this issue but I do wish to mention that the guidelines developed by TESOL for the preparation of ESL teachers are especially worth reviewing and taking into consideration. In our own Consent Decree guidelines we adopted one in particular which suggests that teachers of ESL should "have had the experience of learning another language and acquiring a knowledge of its structure; and have a conscious perception of another cultural system." Sonia Rivera, Vice President of NYSABE paraphrases this thought most admirably in a report on teaching of native Spanish speakers by stating "If you're not bilingual and bicultural by birth at least try to be so by heart."

The training of teachers, of course, depends heavily on what it is that is to be taught and how it is to be taught. In short, the curriculum and the methodology of ESL must also be considered. In this regard, perhaps we cannot dismiss too quickly Dr. Troike's concern with the state of the art of ESL. It never hurts us to periodically assess our teaching practices and to modify and strengthen our methodology and our instructional materials. We can avoid criticisms that ESL programs are rigid or narrow and limited to pattern drills by continually evaluating and improving our work so that ESL can be perceived as a practical, experience-based comprehensive second language arts program.

I am especially concerned with this issue because of its impact on bilingual education. In today's climate of demands for a return to basic skills and minimum standards of competency, the eyes of the public are focused sharply on the progress of our bilingual programs. Justifiably or not, the aspect of our effort which receives the greatest attention is the students' achievement in English. It is essential to note therefore that the success of many bilingual programs will be measured by how successful the ESL instruction has been. I am, therefore, extremely serious when I say that we are counting on you to provide the most effective and highest quality ESL programs you can deliver.

As a member of the Executive Board of the National Association for Bilingual Education and as its immediate past president, I am confident that we are ready to present a united front together with TESOL to provide the leadership necessary to guide the public and the profession out of the confusion and controversy in which others want us to stagnate. Government and other agencies must be made to understand that NABE and TESOL will continue to expend every effort necessary to push for the development and support of the kinds of programs we know are best for our students. The efforts of our friends in the legal profession for the achievement of equal education opportunity have been admirable indeed, but we cannot surrender our obligation to take a forceful hand in determining the goals of education for this society. In short, we must counsel the counsellors, educate the educators, congress with Congressmen, and communicate with the community.

I would like to return to my biological analogy to say that ESL and bilingual education are well and growing stronger because of our healthy symbiotic relationship. If the point needs further confirmation, I would like to announce that at an executive board meeting held recently in our New Orleans conference, the National Association for Bilingual Education officially endorsed the TESOL "Position Paper on the Role of English as a Second Language in Bilingual Education." Our appreciation goes to the committee who did such an excellent job in preparing the statement, to its Chairman, Carmen A. Perez, to Jim Alatis, to Christina Bratt Paulstors, to the TESOL Executive Committee and to all of you, our colleagues. I pledge to you the very same support you have demonstrated for us through this position paper and through so many other TESOL activities.

To conclude, Maria Swanson, new President of NABE invites you all to recuperate in Puerto Rico next year at the NABE conference after you've returned from your hard work in Mexico City. Hasta entonces, adios.

Cross-Cultural Interference in Black-White Relations: Discord in the Classroom

Donald M. Henderson

University of Pittsburgh

This paper is about discordance. It is argued here that the relationships between black children and the schools they attend can be characterized as discordant. Discord results from cultural conflict in the school, and influences the nature of the adjustment between children and the schools. These adjustments have been observed by many as often being considerably less than desirable and appropriate. Both Carter G. Woodson and Melville Herskovits discussed the mis-education of blacks in the latter 1930's. Almost twenty years later, the rejection of the separate-but-equal system of education by the U.S. Supreme Court not only outlawed segregated education, but also called attention to the importance of certain social, cultural, emotional and psychological factors in the child's interaction with the school. In the sixties, these factors became important components of theories of cultural and social disadvantage, which were used to explain why the school performances of black children were so different (much poorer) than those of their white counterparts. These theories were also the bases for a large number of Federal programs such as Headstart, Follow-Through, Upward Bound, etc.,—all aimed at improving the performance of black children in school situations. The proponents of these notions rejected previously popular theories that held the cause of such poor performance to be a function of the genetic inferiority of black children. These theories assert that black children began school with disadvantages because they generally fail to have some optimal quantity of certain social and cultural experiences that are presumably vital to success in school. Other academically successful children apparently enjoy these vital experiences in the normal course of becoming school-aged and, therefore, acquire those experiences that prepare them for success in school. Black children, however, are perceived as suffering from a variety of deprivations, ranging from intellectual to financial. Each deprivation is seen to have significant impact on the children's preparedness for and ability to function appropriately in the school house. Labov's characterization of the concept of verbal deprivation is illustrative of this. Noting that verbal deprivation is one facet of the more general notion of cultural deficit, he observes that deprivation theorists accept that:

> . . . Negro children from the ghetto area receive little verbal stimulation, are said to hear very little well-formed language, and as a result are impoverished in their

> means of verbal expression: they cannot speak complete sentences, do not know the names of common objects, cannot form concepts or convey logical thoughts. (Labov, 1975, p. 90)

Labov rejects the notion of verbal deprivation and observes that the "concept of verbal deprivation has no basis in reality". He also notes that black children receive a great deal of verbal stimulation and possesses the same skills with the language as do others who learn to speak a language. He notes that the concept of verbal deprivation is:

> . . . particularly dangerous, because it diverts attention from real defects of our educational system to imaginary defects of the child; (which leads) its sponsors inevitably to the hypothesis of the genetic inferiority of Negro children which it was originally designed to avoid. (1975, p. 90)

Most of the educational programs designed to up-grade the educational performances of black youngsters begin, as Labov observes, with the acceptance of defects in the child that are occasioned by defects in his/her family and community. The programs then are designed to offset these influences.

Just as Labov observes that no such thing as a verbally deprived child exists in reality, I suggest that culturally deprived children do not exist in reality either. If one views the black population as a separate and distinct ethnic community, what is observed as indicators of deprivation from one point of view can now be construed as indicators of the presence of cultural differences. Therefore, children raised in the black community learn particular sets of values and norms, inculcate these orientations into their social selfs, become familiar with particular institutionalized role responsibilities and privileges, share certain beliefs and attitudes, and think and act in ways that are consistent with the culture of the black community. In short, they learn to be Afro-Americans. The values, norms, attitudes and ways of acting, i.e. culture, thus learned enable the child to interact appropriately, that is to say, normatively, within the community of orientation.

Robert Blauner (1970) notes that black ethnicity, although different in its history and development from that of other hyphenated ethnic groups, provides black Americans with a perspective on life that is uniquely black and not shared by other groups. He suggests that Afro-American culture is characterized by its own "distinctive traits and ethos". He notes also that in characterizations of Jews, Italians and Greeks as ethnic groups we implicitly have in mind differences in culture. The same observations hold for Afro-Americans. ". . . Ethnic subcultures reside in a certain number of distinctive values, orientations to life and experiences and shared memories that co-exist within the framework of the general American life-style and allegiances . . ." (Blauner 1970, p. 350) The ethnicity of Afro-Americans is rooted in the common historical and social experiences of the group. Such common experiences can be seen to be reflected in collective social practices and styles of life. The common origin of the collective ethnicity is the African cultural and social heritage, which formed the basis for the adjustments made by the African slaves to their new life circumstances in America. The Afro-American community, with its distinct ethnic character, produces thereby, a distinct group of Americans. They are distinguishable by means of their unique cultural background and social styles. They identify themselves as a distinct group and are perceived this way by others. In short, Afro-Americans are as culturally different in their own days from mainstream Americans (white Americans) as are Chinese Americans.

Growing up black in America means growing up differently from other Americans in a variety of ways. The same general observation can be made about any other ethnic group. The particular differences that characterize the many ethnic groups result from their origins on the one hand, and their historical and contemporary experiences in America on the other. Black children are in fact not ready for the school when they arrive at its door; nor is or has the school been ready for them. Their unreadiness is not a function of cultural or social deprivation, but rather, it is a function of cultural and social differences. The children are only observed to be culturally and socially deprived in the school or in similar institutional settings. The same children, who, as a result of their contact with the school and its functionaries, are classified as slow, dull, incapable of learning, or emotionally disturbed and so on, can be observed in their home communities behaving in ways that suggest they are quick, alert, bright, stable and capable of learning. The children are not ready for the culture of the school and roles they are expected to play there. At age four or five very few children are ready for sustained and intense de-socialization and re-socialization outside their community of orientation.

The public school is often a place of de-socialization and re-socialization for a good many black children. It's possible that this is less the case in traditionally black schools with black faculty and staffs, although desegregation may be changing that. Despite its presence in the black community, the school is more often than not in the community, but not of the community. The school reflects the culture of the wider society[1] and, as such, is itself unready for the presence of the culturally different black child. Mass education assumes that the students are characterized by common backgrounds, common life styles, common beliefs and ways of ordering reality. For most Americans this assumption is functional, but for a minority it is dysfunctional or discordant.

Dell Hymes has observed this among Indian children in the Southwest, whose silence in the classroom puzzles their teachers.

> Again, a child raised in a Cherokee community . . . in some sense knows that certain kinds of conduct show respect for others, regard for other's personal worth, and that other kinds of conduct show disrespect and are shaming. The child may not have brought such patterns to consciousness and may not have a terminology and frame of reference in which to analyze them closely; neither may the child's teachers; yet seemingly small details may make a classroom a subtle battleground of silence . . . (Hymes, 1972, p. xvi)

Hymes notes that "inadequate mutual comprehension" of conduct on the parts of both teachers and students leads to situations that frustrate both of them, and decrease the likelihood that educational efforts will be successful.

---

[1] This is a loose reference to the collectivity of non-minority Americans. America is a pluralistic society.

Hymes' notion of "inadequate mutual comprehension" calls attention to the fact that the teacher and the child both have inadequate comprehension of the nature of each other. The child, at least early in his/her school career, is likely to expect that adults in the school will act more or less like adults do at home and in his/her neighborhood. This especially applies to black adults—while expectations of white teachers will include the child's understanding of the community's ideas about relations between white and black people. The teacher, on the other hand expects the child to come to school and act like other—that is to say, white—students do. The expectations of both are frustrated in a variety of ways, some subtle and some blatant. Attempts on the parts of both to come to terms with each other are constantly frustrated or at best sporadically successful. The understandings of various social performances in the school are imbedded in two separate cultural systems—hence, the inadequate mutual comprehensions of various modes of conduct or, stated differently, discordance in the classroom.

The child, as a product of the Afro-American community, has developed: 1) a fund of cultural competencies, including knowledge of the "shoulds" and "oughts", the "goods" and the "bads", and how to show deference, display respect and to recognize respect, etc., when it is conferred upon him/her, among other things; 2) a repertoire of social performances, based upon understanding of and experience with the institutionalized role sets of the community of orientation; and, 3) the development of a personal way of expressing the oughts etc. through social performances, both of which are importantly influenced by significant others. The school settings, institutions of the wider society, are not accepting of nor adaptive to these particular qualities of black children. They are, therefore, environments of cultural discord and social conflict. For example, the animosity between black and white high school students can be seen as a product of cultural conflict.

What the deprivation theorists observed as indicators of cultural and social deprivation can now be seen as a product of the interaction of the culturally different child and the school. If the school situation is not adaptive to the different styles of life and styles of learning, the different forms of competence and the different modes of performance, the child is caught up in a depriving environment. Hymes speaks of linguistic deprivation from this point of view:

> Children may indeed be "linguistically deprived" if the language of their natural competence is not that of the school; if the contexts that elicit or permit use of that competence are absent in the school; if the purposes to which they put language, and the ways in which they do so, are absent or prohibited in the school. The situation of the children, indeed, is much worse than "deprivation" if their normal competence is punished in the school . . . Such is the fate of many children in classrooms today if the accent, grammatical characteristics, or style of their normal community are heard. (Hymes, 1972, p. xx-xxi)

He characterizes the school and its educational program as repressive. Roger Shuy found that Detroit urban teachers knew very little about the nature of the language spoken by their students and even less about what to do about it in the classroom. Shuy (1975) and others conclude that teachers are not being adequately trained for successful interaction with their black charges.[6]

It's worth noting that discord may, in part, occasion the development of what has been referred to elsewhere as *kiddie kulture*. In the presence of intense discordant situations the children work out ways of adapting to the differential demands of the two competing cultural systems which are different from both of them. These adaptive responses are shared by other students and passed on from one student generation to another. In a more general way, the school in the Afro-American community can be viewed as entertaining three cultural sub-systems.

In summary, children from black communities and the schools they attend are often out of phase with each other. The classroom is a reflection of the values, norms, attitudes, beliefs, expectations and aspirations of the wider society, while the children reflect the more parochial values, norms, attitudes, beliefs, expectations and aspirations of their community of orientation. The relationships in which the two are involved are often mutually non-compatible because of these differences, and hence here characterized as discordant. Hymes refers to the classroom as a "battle-field of contention between conflicting conceptions of such things as values, norms, etc.". In instances too numerous to count, where there are conflicting conceptions of what constitutes appropriate behavior, the situation in which the conflict occurs produces discord. In a broader view, discord can be seen to be one form of cultural conflict. The discordant character of the relationship is further compounded by the fact that both the child and the teacher are often unaware of the cultural gulf that separates and frustrates them. Because both share the dominant themes of the society, it is assumed that the ways in which they experience these themes are more or less the same. It is clear that sub-cultural patterns, articulated within the broad parameters of the dominant societal themes, are the realities that organize and govern the social lives of their populations. On the street, people refer to this phenomena as "different strokes for different folks".

Christina Paulston reports on her experiences after visiting Stockholm, her home town, for the first time in many years. She noticed that although she had retained her fluency with the Swedish language, certain "aspects of Swedish culture had become foreign (discordant) and she consequently failed at times to communicate efficiently. As a result, she was engaged in a series of discordant relationships. She reports several examples of such events. What is common to all of these situations is the fact that without thinking about it she applied American cultural orientations to activate her own social performances in a Swedish (non-American) setting. The relevant cultural and social standards, against which her behavior and that of her relatives and friends was to be judged were Swedish. The orientation undergirding Christina's behavior was not Swedish and was, in this situation, in violation of the relevant norms and customs of Swedish culture. Her behavior, in short, was inconsistent with what "was expected" of a Swede. Had she been American her friends and relatives might have been much more tolerant of her American behavior. After reflecting on this experience she makes the following observations:

> Clearly, the meaning of an interaction is easily misinterpreted if the speakers don't share the same set of rules (as in cases of) non-effective communication where the same surface structure (carries) different social meaning. (Paulston, 1974, p. 347)

Elsewhere she characterizes these situations as follows:

> . . . an encounter where different social meanings are encoded in or decoded from, the same linguistic expression. In such a case we get communicative ambiguity, the same surface behavior but with deep structures in two separate sets of communicative competence. (p. 350)

Like Christina's friends and relatives or like Christina, depending on the point of view you wish to adopt, teachers often incorrectly assume that the "deep structure" underlying the "surface behaviors" of the black children whom they teach is the same as theirs, and, if it isn't, then it's their job to make it so or to try. They, therefore, incorrectly respond to the child's behavior. At the same time, the child makes the same type of incorrect assumption about the teacher by incorrectly responding to the teacher's behavior.

Consider the following encounter in this light. In the black community the child learns that old people are due deference from him/her, and, depending upon their stations in the community and relationship to the child, they are to receive more or less deference. The child also learns the proper social performances for expressing deference and the situations in which they are required. Since social relationships are intrinsically reciprocal, the child also learns how to recognize and interpret adult responses to deferential performances.

One of the ways of conferring deference to older and respected persons in the black community is to avoid looking into their faces and certainly not in their eyes. I can recall, as can other blacks, learning this lesson painfully. It was not uncommon for adults, into whose face the child looks unnecessarily long and/or hard, to construe this behavior as a disrespectful and sometimes, a belligerent act and to respond to the perceived disrespect and belligerence. In the white community not looking into the face of the person with whom one is interacting is often construed as shifty, suspicious and/or anxiety-ridden behavior. Eye-contact avoidance is perceived as an indicator of dishonesty or dislike. As we've seen, the school and often its functionaries reflect the culture of the white society. White teachers, and some few black teachers, have expressed to me their exasperation with their black students who won't look them in the eye when speaking to them. One teacher told me she sometimes just wanted to "shake them until they did". Here we have an instance of inadequate mutual comprehension. The surface behavior (social performance) is comprehended in accordance with two very different "deep structures" (cultural competencies). The teacher and the students are involved in a very subtle and unrecognized battle for cultural survival and supremacy. This battle is seldom resolved in ways that importantly permit the children to gain positive reinforcement for their community selves. However, without such positive reinforcement it is unnecessarily difficult, and too frequently, impossible for the child to develop an adequate school self. The differential requirements placed upon the child by the two communities make the school and especially the classroom, a very confusing and frequently painful place to be. The child is "damned if he does and damned if he doesn't." The child doesn't know the rules by which the school game is played and it's likely that black parents don't either. On the other hand the teacher, if white, quite likely, has another version of the rules of the game, and if black, probably doesn't recognize the problem in a formal way. The child and the teacher comprehend each others' school performances through two different and often mutually incompatible cultural orientations (deep structures). The child thinks the behavior is good and the teacher thinks it's not; so, they're in trouble from the start.

Let's go back to the eye-avoidance behavior. Like any other social performance, it is a part of an intricate web of associated performances, each rooted in an aspect of cultural orientation. Because black children have internalized eye-contact-avoidance they engage in this behavior automatically, which means that the teacher may rarely catch them looking at him/her. I've heard teachers observe that their black students never pay attention in class. It occurred to me that they may seldom be able to observe black children looking directly at them. I used to "crack-up" watching my daughter and her friends imitate their teachers. They corrected each other, made suggestions for improving the imitation; in short, they obviously had done a lot of direct-looking at their teachers—but they hadn't been caught! And, because they hadn't, they were perhaps perceived as being inattentive, when all they might really have been was deferential. *That's the essence of the discord.*

There is, however, a way of indicating that one is paying attention, widely practiced in the black community, but not in the school. One does so by *saying* something to the speaker which suggests, "I'm paying attention to you!" This is institutionalized in that the speaker expects this type of response. Hence, the interaction is importantly tied to these mutually comprehended expectations. In the classroom, spontaneous verbal affirmations of this type are not permitted. Earlier I suggested that the school is often *in* the community, but not *of* the community. It's clear that "going to school" is in a great many important ways discontinuous with growing up Afro-American.

Eye-contact-avoidance behavior is a subtle example of discord in the classroom. The consequences of such discord are likewise subtle and difficult to discern for people not formally aware of the character of two different "deep structures" underlying the meanings read into the same social performance. Another less subtle discordant encounter in the classroom is much more directly reflective of the role of discord in learning itself. Several years ago, a systematic attempt to enhance the verbal capacities of black children in a large urban school district was spectacularly unsuccessful. Black children in this school system were observed to suffer from verbal deprivation.

An elaborate program was devised for increasing the verbal capacities of the children. One of the important learning, facilitating and reinforcing devices of the program was rhyming. The developers of the program counted on the native interest of children in rhyming to help make the learning of new word-skills interesting. Unfortunately, no one took into account the fact that the rules for making rhymes in the black community emphasized sound agreement and the school rules emphasized visual agreement; thus the teacher and the students played the same rhyming game by different rules. As a result, both the children and the teachers grew so frustrated that the effort had to be discontinued.

Rhyming in the Afro-American community is accomplished primarily by making the *sounds* of word endings compatible, while in the school it is accomplished primarily through making the *spellings* of word-endings compatible. For example, "rough and stuff" or "tune and soon" or "hand and man" all fit the rhyming requirements of Afro-American practice but all violate the rhyming requirements of the school practices. In short, the processes used to accomplish the learning goals of the school are sometimes at cross-purposes with themselves. As was the case with the example about eye-contact-avoidance, the surface behavior (rhyming) is produced and interpreted in terms of two different "normative" systems. It is not the case, as the deprivation theorists would have us believe, that the children do not understand how to manipulate the sight-rhyme process, because they do. They don't understand why it isn't also permissible (normative) to use the sound-rhyme process. As Hymes has observed the absence of opportunities for Indian children to use their cultural skills in school is cultural repression. The fact that the child is punished for sound rhyming and not for sight rhyming is not only confusing for the child, it is also repressive.

Elizabeth Whatley (1973), in studying the appropriateness of teacher assessments of verbal responses of black elementary school children in Pittsburgh, found the standard teacher assessments to be incorrect when the child's responses were judged using the phonological requirements of black English, as opposed to those of standard English in an alarmingly high proportion of the cases. Of the 100 cases examined over 7 dimensions she found teachers judged the children incorrectly forty-four to seventy-seven percent of the time. The child is involved in a situation where his/her "natural" (culturally stimulated) social performances are systematically stiffled. Over time the child learns to operate within the restricted parameters of sight-rhyming but certainly not without a stifling of initiative and a blunting of creativity. For far too many black youngsters the school is not a place where they can further the social development begun in their communities. Athletic activities may offer one of the few areas, if not the only one, where the skills developed by the child in the community of orientation are accepted. This may, in part, contribute to the excellence achieved by black student-athletes.

I'd like to examine the differential operation of social control in the Afro-American community, on the one hand, and the school, on the other. Some recent work done with black elementary school children in Pittsburgh by Alfonzo Washington and I (Henderson and Washington, 1975) led us to conclude that adults in the Afro-American community perform very different social control roles than do adults in the school. All adults in the child's social space are expected to respond to the child's social behavior in much the same way as his/her parents behave. The child is, therefore, constantly under the surveillance of a network of adult social control agents. Infractions of community norms reprimanded by control agents subsequently are reported to parents via this control network. In short, adults are actively responsible for "controlling" the activities of children. ". . . Adults in the school did not fulfill functions of social control in ways that (are) consistent with the child's expectations of how adults should behave in response to him in situations that require the enforcement of social controls. (Henderson & Washington, 1975, p. 357) The school's control system rarely included the parent, except when a major sanction was invoked, such as suspension or dismissal. The focus of social control is expected to reside in the child—that is to say, the child is expected to police him or herself, and indicates an aspect of readiness for school. In point of fact, however, the Afro-American child is brought up to respond to the external control activities of adult social control agents. The child's expectations are that adults in the school will behave like adults in the community of orientation. Unfortunately, these expectations are not fulfilled. It is our hunch that the breakdown of discipline in schools in the black communities in large measure results from discord in the control system.

Discord in the classroom constitutes a problem of major importance in the education of black children, as well as children from other distinguishable ethnic minorities. It is imperative that these problems be squarely faced in our attempts to educate black youngsters. It is certainly possible to make the adjustment of black children to the school an easier thing to realize. I have seen it work. For example, when my son, Mark, was in the third grade, his class learned to sing, "Go Tell It On The Mountain" incorrectly. After hearing him sing it incorrectly, I taught him "how to do it right" and told him it was a song made up by black folks. I also encouraged him to teach his teacher how to sing it "right" so she could teach the class the correct version. The following day he related how he had mentioned what I said to his teacher, who asked him to teach the class the correct version. He felt very good about an encounter which was initially discordant. His teacher had acknowledged his ethnicity and demonstrated to the class that he had a positive contribution to make to the class precisely because of it.

And that, after all, is what education in a pluralistic and democratic society is all about. The realization of the promise of the American dream can be importantly furthered through education which authenticates the value of pluralism by living and learning about each other.

REFERENCES

Blauner, Robert. 1977. Black culture: myth or reality. In Norman Whitten and John Szwed (eds.), *Afro-American anthropology*, Glencoe, Ill., The Free Press.

Henderson, Donald M. and Alonzo G. Washington. 1975. Cultural differences and the education of Black children: an alternate model for program development. *Journal of Negro Education 54*, 3, 353–360.

Hymes, Dell. 1972. Introduction to *Functions of language in the classroom*, C. Cazden, V. P. John, and D. Hymes (eds.), New York, Teachers College Press, Columbia University.

Labov, William. 1975. The Logic of Nonstandard English. In Paul Stoller (ed.), *Black American English*, New York, Dell Publishing Company (Delta Book).

Paulston, Christina Bratt. 1974. Linguistic and Communicative Competence. *TESOL Quarterly 8*, 4, 347-362.

Shuy, Roger W. 1975. Teacher training and urban language problems. In Paul Stoller (ed.), *Black American English*, New York, Dell Publishing Company (Delta Book).

Whatley, Elizabeth. 1973. The Identification and Phonological and Syntactical Variants in the Oral Reading of Black Dialect Speaking Children. Unpublished dissertation. University of Pittsburgh.

Part 2

Teaching: Theory and Practice

The Preparation of Grammar Explanations in the Teaching of English As A Second Language

Patricia R. Furey

English Language Institute
University of Pittsburgh

The purpose of this paper is to offer some guidelines for the construction and adaptation of grammar explanations in the teaching of English as a Foreign Language. The grammar lesson typically consists of two components: 1) exercises and activities requiring the student to in some way manipulate the structure under consideration and 2) a grammar explanation which explains, or in more inductive approaches, clarifies the structural principles being taught. This latter component of the grammar lesson has, in recent years, been largely neglected in discussions of language learning materials and methods, a situation which would seem to be primarily due to the influence of inductivist audio-lingual methods. The audio-lingualists warned us to teach the language, not to teach about the language. To their way of thinking, the grammar explanation serves merely as a rule clarification or summary, the purpose of which is to confirm generalizations previously reached by the student during extensive oral drill sequences.

For many of our students, however, the grammar explanation may play a more crucial rôle than that attributed to it in audio-lingual approaches, and while we don't want to turn our language lessons into lengthy linguistic analyses, there are several arguments which support the position that more careful attention and increased emphasis be given to the presentation of grammar rules.

First, many of our students, particularly adolescent and adult learners, seem to want more grammar explanation. Teachers frequently comment that their students, especially those beyond the very initial stages of instruction, ask for more discussion of the rules underlying the structures they are learning. Many individuals want and seem to need to know more about how the language is put together.

The desire of at least some of our students to have more rule explanation may indicate something about differences in learning strategies. Some individuals may learn more effectively through deductive strategies, requiring understanding of general principles prior to their application in language activities and exercises, and thorough, carefully constructed grammar explanations would seem to benefit this type of learner. Furthermore, during the course of a typical grammar lesson, our students are assaulted with a great deal of oral language.

Seeing the structures under consideration within the context of the grammar explanation provides for some learning to take place through the visual modality, a fact which is of particular advantage to our visual learners.

The need for thorough, efficient rule exposition in the language lesson is recognized by such authorities as Rivers (1970) who stresses the importance of full understanding of rules, particularly those relating to the meaning or function of structural patterns, and Carroll (1974) who advises the textbook writer to provide descriptions and analyses of how the language is put together. Furthermore, the literature in educational psychology and learning theory advises us that understanding of underlying principles assists learning (e.g. Cronbach, 1963; Glaser, 1966; Ausubel and Robinson, 1969). The advantages of carefully formulated rule presentation can best be understood with reference to the concept of meaningful learning which, according to Ausubel and Robinson (1969), results when the learner attempts to relate in a sensible way the material being learned to what he already knows. It is in the grammar explanation component of the language lesson that we can substantially assist this assimilation and integration of newly learned material into the student's cognitive structure, and specific ways of facilitating meaningful learning through grammar rule presentation will be explored later in this paper.

The construction of rule presentations which facilitate functional control of structural patterns requires first a clear understanding of exactly what kind of phenomenon the grammar explanation is and precisely what it contributes to the language learning process. Paulston and Bruder (1975) identify three component parts of the grammar explanation. The formal analysis is that part of the rule presentation which demonstrates the form of the pattern under consideration. It makes clear the linguistic elements which are present in a given pattern and indicates how they are arranged and related. The statement "We add *-er* to most one-syllable adjectives to form the comparative", for example, is part of the formal analysis in an explanation of the comparative pattern.

In addition to knowing the correct form and arrangement of a given structural pattern, the student must also understand how to use that pattern. That is, he must be aware of how it conveys meaning so that he will be able to select it from all of the structural possibilities of the language in order to efficiently and accurately express his intentions, thoughts, feelings, etc. Criteria for selection of patterns are provided either explicitly or implicitly, in that part of the grammar explanation known as the functional explication. The functional explication gives those conditions which must prevail in the "real world" at large, and in the context of the communicative situation in particular in order for certain forms to be selected. The statement, *We use the simple present tense to indicate a habitual action or state*, is an example of functional explication.

In the case of many patterns, the functional explication is the most difficult part of the grammar explanation to formulate due to the difficulty of stating with any degree of precision and pedagogically useful generalize-ability the criteria calling forth the use of certain structures. The conditions surrounding the use of a given pattern may involve highly complex and subtle considerations of what has been said previously in the conversation, of complicated time relationships, and of the assumptions and presuppositions of both the speaker and listener. The definite article in English, for example, is one pattern for which the functional criteria for usage are extremely complicated.

In addition to formal analysis and functional explication, lexical clarification is required in the rule presentations of certain patterns. Lexical clarification, which is essentially an aspect of functional explication since it indicates meaning, is applied primarily to patterns within which lexical items from a closed set may fill a given slot. The student must know the lexical meaning of these alternate forms if he is to use the patterns correctly. Lexical clarification is necessary for forms such as indefinite pronouns, modal verbs, frequency adverbs, and expressions of quantity. The statement "The modal verb *can* indicates ability" is an example of lexical clarification.

With an understanding of what elements constitute the grammar explanation we may now proceed to some guidelines for its construction. While most directly addressed to those designing their own grammar explanations, these guidelines can be readily applied to those adapting explanations from existing materials. Also, it should be stated that the guidelines are offered without explicit reference to any particular theory of grammar.

In designing a grammar explanation the first step is to locate thorough and accurate linguistic information about the pattern under consideration. Grammar texts, reference grammars, and articles concerned with ways of presenting particular structures in journals such as *The TESOL Quarterly* or *English Language Teaching* are all sources which provide an information base for the construction of the explanation. An important point here is to examine carefully several sources to assure that all the relevant information about the pattern is obtained.

After careful examination of the linguistic facts pertaining to the use of a given pattern, we must select that information to be presented in the explanation. This is one of the most crucial and difficult steps not only in the construction of the grammar explanation but also in the design of the entire grammar lesson for it is here that we are confronted with the problem of how much information to present at a given stage of instruction. Certainly at the beginning level we don't want to hinder learning or cause motivational problems by overwhelming our students with more information than they can assimilate at a point when they have had relatively little experience with the language. Yet even at the beginning stages of instruction there are patterns which by virtue of their high frequency must be taught even though the rules governing their use may be quite complex and detailed. For these patterns, such as the definite article or the present perfect tense, we may have to settle for presentation of a limited number of rules, particularly functional rules, at the beginning levels. In commenting on how much to include in the grammar explanations of beginning level texts Bolinger (1968, p. 27) states: "The writers of textbooks must always face the problem of how explicit their rules are to be. A partial formulation (which is inadequate by definition) is as often as not a requirement, in a beginning text, to launch the student into the production of sentences of a limited range, in order to have him produce anything at all . . ."

In an interesting and useful article calling for more rational criteria in the selection of material to be taught, Johnson (1969) notes that in deciding what and how much is to be presented in the language lesson we must consider the time taken to develop control of a form and weigh it against the relative value or importance of that form in the total communication process. Thus, for example, it hardly seems worthwhile at the initial stages of instruction to expand very much time and effort teaching the numerous conditions for use of the definite article since it is such a difficult form for most learners.

But while we must be careful to select and elaborate our teaching points in consideration of the relative benefits derived from the amount of time spent on them, we must not forget that at a certain stage our learners will be frustrated if they have not been given enough rules to enable them to use the language fluently and appropriately. It is here that the concept of spiraling can be applied to the design and sequencing of our grammar lessons in general and to our rule presentations within these lessons. Ausubel and Robinson (1969: 318) define the spiral kind of organization as one in which the same topics are ". . . treated at progressively higher levels or sophistication in successive sections." We might apply the concept of spiraling in teaching the present perfect tense for example.

Crowell, in his *Index to Modern English* (1964, p. 388), describes three functions of the present perfect tense: 1) to indicate an action or state which began in the past and continues up to and into the present; 2) to indicate an action or state repeated in the past and which will possibly be repeated in the future; and 3) to indicate an action that occurred at an unspecified time in the past. The first functional rule for the present perfect might be presented, along with the formal analysis, at a relatively early point in the instructional sequence; the other functional rules should then be taught when the present perfect is re-introduced at a later time.

Another point to be considered in relation to the issue of what and how much information to include in the grammar explanation is the presentation of some rules for the purpose of pattern recognition rather than pattern production. For example, the simple present tense is sometimes used to indicate future time especially with verbs of motion, such as in the sentence *The plane leaves at 10:00 tonight.* (Crowell, 1964, p. 387). Students do not need to produce such sentences since there are several other patterns which more commonly indicate future time. However, a note in the grammar explanation referring to this rule will help the student recognize and understand this use of the simple present tense when he hears it spoken or sees it written.

Certainly at the more advanced stages of instruction we need to include more information in our explanations than at the beginning levels. The advanced patterns of English are formally and functionally more complex, and although we don't know precisely what makes a given pattern difficult to learn, it is clear that generally the material at the advanced level is harder to teach, harder to learn, and simply requires more intellectual analysis on the part of the student. Relative clauses, comparatives, superlatives, and the perfect and perfect continuous modals provide examples of patterns which testify to the necessity of including a good deal of linguistic information in our advanced level explanations.

In deciding what and how much information to include in the grammar explanation then we must consider the level of instruction and the difficulty and complexity of the patterns under consideration. We can benefit from information provided by research in such areas as contrastive analysis and error analysis which point out possible sources of trouble for students and thus indicate what we might include in our rule presentations. But we must also draw largely on our previous teaching experience and our pedagogical intuitions about how much and what kinds of information the student can meaningfully assimilate at a given point in the instructional sequence.

There are, however, several firm rules to keep in mind when selecting material to be included in the grammar explanation. First, we must be sure the explanation is well integrated with the overall grammar lesson by including all the linguistic information necessary to enable the student to complete any associated activities and exercises.

Second, we must be certain that the kind of information included in the rule presentation is consistent with the overall purpose of the materials. It is important to keep in mind the skill we are teaching and to include the appropriate information. Thus, the grammar explanation included in *writing* materials may need to incorporate certain punctuation rules, while the explanation within materials designed to teach speaking should include any pronunciation rules which pertain to the use of given grammar patterns, e.g. the various pronunciations of the third person singular present tense endings.

Also, in determining what information to put in our rule presentations we must remember to provide any necessary reference lists or items associated with a given pattern. In the presentation of the present perfect tense, for instance, we would want to include a list of the irregular past participles, or in an explanation of the gerund we may need a list of verbs which are followed by the gerund as opposed to the infinitive form. In order to assist efficient learning it is important to arrange our listed items according to any relevant organizing principles. Thus, for instance, we would group together those irregular past tense verbs whose forms are similar (e.g. teach/taught, think/thought, bring/brought, etc.).

Finally, in considering what information to include in our explanations, we should give some thought to incorporating communicative competence rules for certain patterns. Students need to know when it is appropriate or inappropriate to use certain structures, what patterns are more or less formal, etc. The request form patterns, for instance, call for the inclusion of some of the social rules of language use in the grammar explanation.

Once it has been determined what information is to be incorporated in the rule presentation, we must decide on the best modes of presentation to use in conveying this information. There are four general modes of presentation of grammatical information: verbal generalization, example, iconic device, and symbolic notation. In determining the most efficacious modes of presentation for a given grammar rule, there are two general concepts to be considered. One is economy, and the other is meaningful learning.

According to Bruner (1967: 45) economy relates to ". . . the amount of information that must be held in mind and processed to achieve comprehension." In representing knowledge we achieve economy by simplifying terms in certain ways. Some modes of presentation for a given quantity of information are "simpler" than others; thus, economy is largely dependent upon the mode of presentation used in conveying this information. But a note of caution is needed here, for while we want to increase economy, we don't want to sacrifice what Bruner refers to as the effective power or generative value of our rule presentations. In other words, we don't want to simplify our rules to the extent that they are not thorough enough to enable the student to master the pattern under consideration.

The second concept to be applied in the selection of mode of presentation of grammar information is that of meaningful learning. Two of Ausubel's and Robinson's (1969) prerequisites for the occurrence of meaningful learning are: 1) that the material to be learned be logically and systematically organized so that it is capable of being related to existing cognitive structure, and 2) that the learner possess the relevant anchoring ideas in his cognitive network to which the new material can be related. Through careful selection and formulation of the mode of presentation of linguistic information we can facilitate meaningful learning by demonstrating the logical, systematic organization of language and by introducing and reinforcing the appropriate anchoring ideas to which new material can be related.

Bearing in mind the notions of *economy* and *meaningful learning* we can now proceed to more detailed examination of the four modes of presentation of grammar rules. Verbal generalization is one of the most common ways of expressing a grammar principle. *Frequency adverbs follow the verb BE; The simple past tense indicates an action completed in the past;* and *The modal verb* MIGHT *indicates possibility* are all examples of verbal generalizations which indicate form, function, and lexical meaning respectively.

In designing grammar explanations it is important, of course, to formulate accurate generalizations. It is also important to know what not to express by this mode, since there are some patterns for which it may be difficult to simply and clearly formulate the underlying principles, particularly the functional rules. The negative and interrogative forms of verb patterns, for instance, do not require verbal statements of meaning or function—clear examples serve more efficiently to convey the meaning of these structures.

When we do use generalizations in presenting rules, we must be sure that they are clear, concise, and stated in simple enough terms for our students to understand. This is particularly true when using the target language to express the grammar rules, as is often the case in EFL classes. At the beginning levels in particular the use of simple sentence structures, such as the imperative, familiar vocabulary, and common non-technical linguistic terminology in our generalizations contributes to the economy of our explanations and facilitates efficient learning.

A verbal generalization within a grammar explanation generally expresses a principle and is thus a statement of the relationship among concepts. The educational psychologists (e.g. De Cecco and Crawford, 1974; Gagné, 1965) tell us that in order to learn a principle, a student must first understand its component concepts. As a statement of a grammatical principle, then, a verbal generalization indicates what concepts must be understood by the student before he can master the principle, and in this way suggests what specific points of information may need to be either reviewed or introduced in the rest of the explanation. Thus, for example, in an explanation of the quantity words, *much and many*, we might write a generalization such as "*Many* is used with count nouns, and *much* is used with non-count nouns." The concepts of count and non-count nouns might be reviewed if they have been introduced previously, and, of course, would be explained fully if they have not been presented before.

Particularly for some of the more complex, advanced patterns of English it may be necessary to state several formal or functional generalizations within one explanation. The various formal rules pertaining to comparative constructions are one example of where several generalizations may be necessary. In these cases we should make our explanations as economical as possible by stating the rules in the most general terms feasible, presenting the most general rules first, and clearly separating these from exceptions and irregularities.

Verbal generalizations lend the quality of meaningfulness to the learning task by informing the student of the structural principles underlying the language. One way in which our verbal rule statements can be made more meaningful is by using them to demonstrate similarities among different patterns. For example, in showing question formation with modal verbs, it can be stated as in sentences with the verb BE, questions are formed by inverting the subject and the verb. Such statements of similarities among patterns help the student to view the language in terms of its central organizing principles thus enhancing meaningfulness.

Like the verbal generalization, the use of example is a common mode of presentation of information within the grammar text. In most contemporary materials, the example is framed in the form of a sentence illustrating the pattern under consideration. There are two ways of using examples within the grammar explanation. They may be given to support or illustrate an accompanying generalization, or they may be used alone as a basis for induction of a grammar rule, particularly when it is difficult or awkward to formulate the rule in terms of a verbal statement. In either case, there are several points to remember in providing clear, economical, and meaningful examples. First, it is important that our examples illustrate all of the salient formal and functional aspects of the rule at hand. This often necessitates inclusion of several example sentences. Thus, for instance, in presenting the rule that we use the present continuous tense to express an action taking place in the present or near future we would need two example sentences such as: *I am eating dinner now. I am doing my homework after dinner.* It is particularly important to provide an ample number of example sentences with intermediate and advanced level patterns which may be more formally and functionally complex. For instance, when teaching relative clauses it is important to include examples where the clause appears both in the middle of the sentence and at the end of the sentence.

We also want our examples to clearly delineate the principles underlying the pattern. This is especially important when no verbal generalization is given thus necessitating the students' inducing all the relevant rules from the examples presented. For instance, we must be particularly careful to construct our examples so they provide the semantic links necessary to the learners' understanding of the functional conditions underlying the use of the pattern. Thus, for instance, in giving example sentences of verb patterns, we need to include time words which point up the meaning of the patterns. *I finished the work yesterday* is a better example sentence for the simple past tense than *I finished the work*.

Finally, our example sentences should sound natural and should contain only familiar vocabulary.

Another less common but often effective mode of rule presentation is the iconic device which Bruner (1967, p. 44) describes as ". . . a set of summary images or graphics that stand for a concept without defining it fully." Iconic representation is particularly useful in teaching grammar rules to children, but it can also be used effectively with adult learners, especially to demonstrate meaning or function. Time lines, for example, can be used to represent the functional conditions underlying the use of verb patterns. Graphic representation might also be used to provide lexical clarification of the various quantity words. By shading in a circle or bar to represent various degrees of quantity and labelling the shaded areas with the appropriate term such as *a little*, *a lot of*, etc. we can efficiently demonstrate the meanings of these expressions. When used effectively, iconic devices contribute to the economy of the explanation largely by eliminating the need for the student to process verbal statements.

Like iconic representation, symbolic notation as a mode of presentation of grammar information contributes to economy and also enhances the meaningfulness of our explanations. Formulaic expressions such as S + V + O to represent subject-verb-object word order are usually used in conjunction with one of the other modes such as generalization or example. Such symbolic devices contribute to meaningful learning by helping the student to see the formal properties of the language in terms of systematic general principles, and furthermore, do so economically by reducing rules to concise formulas.

Once we have determined the most effective modes of presentation for a given pattern, we must consider how to effectively arrange and organize our information within the explanation. One concept to keep in mind in organizing grammar information is that of contrast. Contrast is a widely used pedagogical device which contributes to meaningful learning by relating the unknown to the known. Within the grammar explanation it is most frequently evidenced in example sentences which contrast the new structure with ones previously taught. Dykstra (1956) notes that contrast is applied most effectively in language teaching when it is manipulated in such a way as to point up *minimal* differences in form and function. Thus, in presenting the simple past tense we might give the sentences:

Tom plays soccer every day.
Tom played soccer yesterday.

By varying only the time expression and the verb forms we can clearly point out both the formal and functional properties of the simple present and simple past tenses.

Another rule to keep in mind in arranging the information of the grammar explanation is to make effective use of perceptual organizers. According to Ausubel and Robinson (1969, p. 317) perceptual organizers provide "built-in mechanical aids that make the material perceptually more salient and apprehensible . . ." They include devices such as bold-face print, underlining, use of capital letters, arrows, boxes, etc. It would be useful to have some empirical research on what perceptual organizers are most effective in facilitating learning. Until such research is forthcoming, we have to rely on our own judgment and experience in selecting and using these mechanical aids in presenting our rules. One firm guideline to keep in mind, however, is to use the same organizing devices consistently throughout the materials, e.g. always use an arrow to indicate subject-verb inversion in question patterns. Another rule of thumb is to use these devices selectively in a way which will clearly point up the most salient formal and functional properties of the pattern under consideration, e.g. when giving a sentence illustrating a particular verb pattern, underline the verb. We must also be careful not to overuse these perceptual organizers, for too many arrows, underlinings, boxes, etc. may only make the explanation difficult to follow.

Finally, there are certain issues in the design of grammar explanations which have not been explored here. Whether to use the native language of the students is one such issue; where the explanation should be located in the sequence of the grammar exercises and activities is another issue. As writers and users of grammar materials, we should examine these questions. We must also keep abreast of relevant developments in both educational psychology and linguistics and be resourceful in applying the findings of these areas to the design of grammar explanations. Work in pragmatics, for example, may reveal improved ways of teaching the function of grammar patterns. In the meantime, we must make fullest use of that information we do have about language and how it is learned in designing grammar explanations which will effectively facilitate mastery of the structural patterns of English.

REFERENCES

Ausubel, David P. 1963. *The psychology of meaningful verbal learning*. New York, Grune and Stratton.

Ausubel, David P. 1968. *Educational psychology: a cognitive view*. New York, Holt, Rinehart and Winston.

Ausubel, David P., and Floyd G. Robinson 1969. *School learning: an introduction to educational psychology*. New York, Holt, Rinehart and Winston.

Bolinger, Dwight. 1968. The theorist and the foreign language teacher. In Harold Allen and Russell Campbell (eds.), *Teaching English as a second language, A Book of Readings*, New York, McGraw-Hill.

Bruder, Mary Newton. 1974. *MMC—Developing communicative competence in English as a second language*. Pittsburgh, University Center for International Studies.

Bruner, Jerome S. 1967. *Toward a theory of instruction*. Cambridge, Mass., Harvard University Press.

Bull, William E., and Enrique Lamadrid. 1971. Our grammar explanations are hurting us. *Modern Language Journal 55*, 449–54.

Call, Mary E. 1976. The importance of grammatical sequencing in materials adaptation and development. Paper presented at the 1976 TESOL Convention, New York City .

Carroll, John B. 1974. Learning theory for the classroom teacher. In Gilbert Jarvis (ed.), *The challenge of communication*, Skokie, Illinois, National Textbook Company.

Cronbach, Lee J. 1963. *Educational psychology*. New York, Harcourt, Brace and World.

DeCecco, John P. and William R. Crawford. 1974. *The psychology of learning and instruction*. Englewood Cliffs, New Jersey, Prentice-Hall.

Dykstra, Gerald. 1956. Perspective on the teacher's use of contrast. *Language Learning 6*, 3 & 4, 1–6.

Gagne, Robert M. 1965. *The conditions of learning*. New York,, Holt, Rinehart and Winston.

Glaser, Robert. 1966. Variables in discovery learning. In Lee S. Shulman and Evan R. Keisler (eds.), *Learning by discovery—A critical appraisal*, Chicago, Rand-McNally and Company.

Johnson, Francis. 1969. The failure of the discipline of linguistics in language teaching. *Language Learning 19*, 3 & 4, 235-44.

Klare, George R., James E. Mackey and Levarl M. Gustafson. 1955. The Relationship of patterning (underlining) to immediate retention and to acceptability of technical material. *Journal of Applied Psychology* 39, 40–42.

Lakoff, Robin. 1969. Transformational grammar and language teaching. *Language Learning 19*, 1 & 2, 117–40.

Lees, Robert B. 1955. Grammatical terminology in teaching English. *Language Learning 5*, 117–21.

Overing, R. L. R. and R. M. W. Travers. 1966. Effect upon transfer of variations training condition. *Journal of Educational Psychology 57*, 179–88.

Paulston, Christina B., and Mary N. Bruder. 1975. *From substitution to substance, a handbook of structural pattern drills*. Rowley, Mass., Newbury House.

Pelosi, Angelo. 1973. What is grammar? *Modern Language Journal* 57, 7, 329–35.

Politzer, Robert L. 1958. Some reflections on the use of the native language in elementary language teaching. *Language Learning 8*, 3 & 4, 49–55.

Politzer, Robert L. 1968. The role and place of the explanation in the pattern drill. *IRAL 6*, 4, 315–31.

Rivers, Wilga. 1970. Rules, patterns and creativity in language learning. In Kenneth Croft (ed.), *Readings on English as a second language*, Cambridge, Mass., Winthrop Publishers.

A Conceptual Framework For Teaching Articles In English[1]

William Acton[2]

University of Michigan

Teaching the use of articles in English to non-native speakers, even if ultimately effective, is often accomplished very inefficiently. Most pedagogical grammars either overwhelm the student with long lists of rules or give practically no guidance at all. Both situations point up a widespread confusion between a linguistic description of a language as opposed to a learner-oriented grammar. For the learner, grammatical rules are only relevant when they assist him in acquiring the use of particular aspects of the language—without getting in his way.

In this paper a schema based on recent and more traditional perspectives on the use of articles is proposed which can 1) assist the learner in understanding why articles work the way they do—especially in terms of the differences between formal and informal usage, and 2) help him predict which article should be used in a given context.

Even though there have been numerous papers and books written on the use of articles in English, most teachers would agree that our understanding of how they should be taught has not been advanced very much since the pioneering work of structuralists such as Christophersen (1939) or Jespersen (1949). That is, with a few notable exceptions, the rules given in pedagogical grammars are typically of the type: "The definite article is used in context X—with the following exceptions: . . ." That rule is *structural* in the sense that it takes a gramatical structure as its point of departure. The problems inherent in a purely structural approach to language teaching have been the subject of a great amount of debate and research.

More recently, a considerable amount of work in theoretical and applied linguistics has been concerned with language use and language functions. A functional, or semantic, approach, on the other hand, begins with a notion such as "definite and indefinite meaning" (Leech and Svartvik, 1975) and then goes on to treat all those forms in the language that can be used to express that function(s). A few good papers and grammars for teachers are available that deal with articles from more or less functional perspectives (e.g. Hok, 1970; Kaluza, 1971; Grannis, 1972; Whitman, 1974; Leech and Svartvik, 1975), but, in general, student grammars and grammar texts continue to reflect essentially structural analyses of the system of articles in English.

---

[1] This is an expanded version of the paper entitled, "A magic rule or two for teaching the use of articles in English" presented at the convention.

[2] I am greatly indebted to Fred Lupke of the University of Michigan for his comments on earlier versions of this paper. In addition Brad Arthur, Debbie Tyma, Ann Borkin and other colleagues at the University of Michigan have contributed on many occasions to the development of the ideas presented here.

The framework to be developed here draws on recent work in language functions (e.g. discourse analysis), as well as more traditional, structurally-based approaches to teaching articles. It has been created with two purposes in mind. First, it is an attempt to interpret certain theoretical insights on articles so that they are directly applicable to the problems of teaching. A second goal is to set up a conceptual framework of article usage, expressed in easily understood English, that will assist the learner by helping him (literally) see the system as a whole.[3]

Part One: Preliminaries

The skeleton of the framework which is presented to students is given in Figure One, A Conceptual Framework for Article Usage. Figure One also reflects the general outline of this paper. The various branches and end points of Figure One will be discussed, proceeding from top to bottom.

Two aspects of this framework, or system, should be emphasized. First, the principle underlying the choice of terms has been basically whether or not a label was readily understood by students. As can be seen in Figure One, some of the terms are from traditional grammars (i.e., proper versus common, abstract, mass), others are not (i.e., bounded versus not bounded, close to Self versus close to Other, new versus shared information). Second, Figure One must be thought of as representing a set of relationships. Students should not be led to believe that the schema "operates" only from left to right, or vice versa. For instance, if one begins at a point where one of the articles actually appears in Figure One (e.g. IIA2 is the location of the definite article), he would have to read in both directions at the same time to find out exactly when it occurs. Again, the fundamental idea here is that, to be maximally effective, the system must be viewed in its entirety, as a whole—not just as a collection of rules. Consequently, there is an intentional de-emphasis on the use of directional statements of the kind: "Use X in this situation" or "In this situation, use X." One assumption behind Figure One, supported to some degree by research and experience in the use of visual aids, is that although such a visual representation may be more or less formally equivalent to a list of rules, to the learner it is potentially a more meaningful and coherent piece of language data.

A basic distinction in this system, located at the far left of Figure One, is

---

[3] "The system" in this paper will, for the most part, be artificially limited to the use of the definite article, the indefinite article and instances where no article is used (the so-called "zero realization" of the article). A complete description of the system of definiteness would go far beyond the scope of this paper. Obviously, grammatical units such as possessives, numbers, demonstrative pronouns, quantifiers and relative clauses have much the same functions as do articles. The functional similarities are not fully developed here, but the slots where those elements would fit in the framework are discussed briefly.

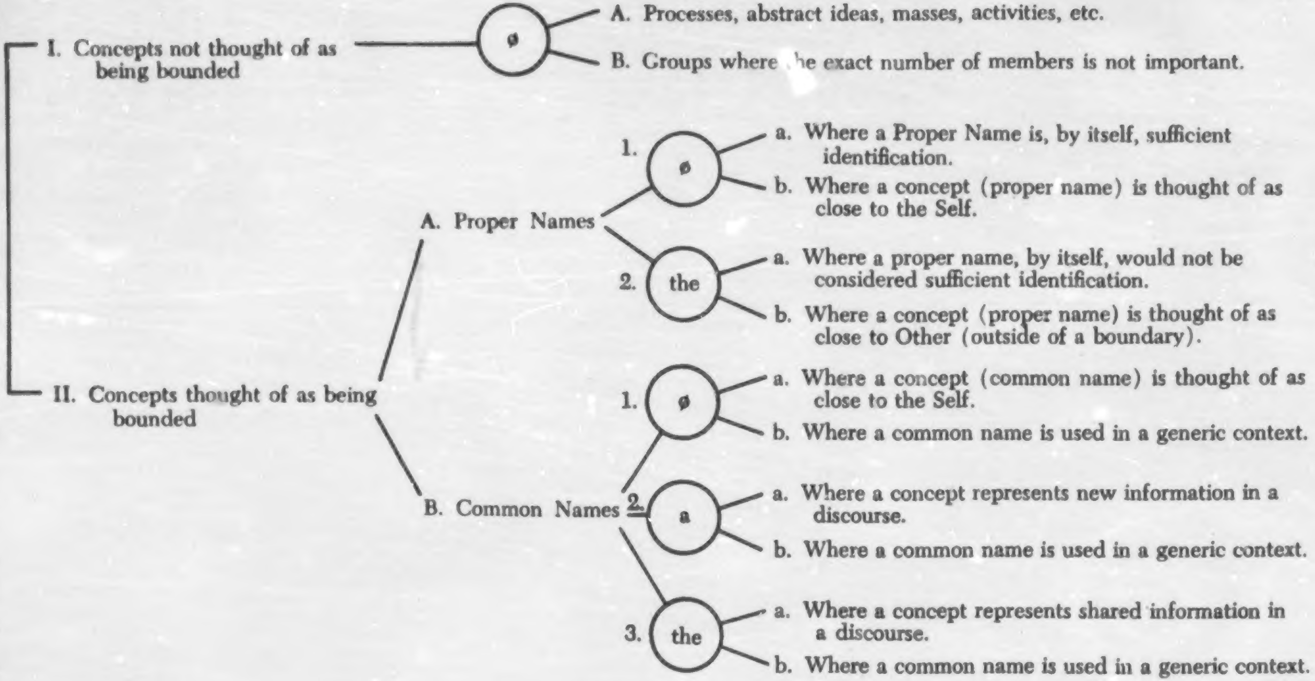

I. Concepts not thought of as being bounded
ø
A. Processes, abstract ideas, masses, activities, etc.
B. Groups where the exact number of members is not important.
II. Concepts thought of as being bounded
A. Proper Names
1. ø
a. Where a Proper Name is, by itself, sufficient identification.
b. Where a concept (proper name) is thought of as close to the Self.
2. the
a. Where a proper name, by itself, would not be considered sufficient identification.
b. Where a concept (proper name) is thought of as close to Other (outside of a boundary).
B. Common Names
1. ø
a. Where a concept (common name) is thought of as close to the Self.
b. Where a common name is used in a generic context.
2. a
a. Where a concept represents new information in a discourse.
b. Where a common name is used in a generic context.
3. the
a. Where a concept represents shared information in a discourse.
b. Where a common name is used in a generic context.

Figure 1. A Conceptual Framework for Article Usage.

the idea of "boundaries".[4] In English, one of the primary functions of articles is to signal that conceptual boundaries have been imposed upon, or placed around, ideas so that they can be treated as objects or units, in some sense. In other words, once a native speaker of English conceives of something as being an object, or having a boundary, he can then count it and/or mentally separate it from something else. The notion of boundaries is similar in certain respects to what has been termed "objectification" (Whorf, 1951) or "serving to concretize" by various linguists. Visual representations of conceptual boundaries (see Figures Two through Four) will be employed here to illustrate for the student (and teacher) four closely related functions of the system of articles:

(1) Articles can make concepts into units (countable)—the traditional count/noncount distinction.
(2) Articles signal whether or not something is to be considered unique or not, whether something can be separated from all like objects and identified mentally—the definite/indefinite distinction.
(3) In a generic construction, articles indicate which type of generic is intended.
(4) Articles are used to indicate whether something is thought of as being close to the Self or more distant from it. (See discussion below.)

Figure Two, A Visual Representation of Bounded and Unbounded Concepts—I (the use of the indefinite article versus no article), demonstrates the first function (1). In Figure Two, the purpose of having the orange juice appear to "flow in" from the right in both A and B is to stress the fact that the boundaries of the orange juice under discussion may or may not be relevant to the conversation. In A, for instance the fact that there could be an infinite amount of the substance off somewhere to the right is a potential reading of the statement "She had orange juice." In B a boundary has been imposed around a quantity of orange juice; the juice has been objectified and can be counted. Likewise, in (5)

(5) She had six juices.

juice has been objectified—just as in the case of Figure Two, B.

The same process is evident in C and D. Here, individual oranges are "rolling in" from the right indicating that the exact number of oranges is in some sense unimportant. "Oranges" in C is treated as an unbounded concept just like "orange juice" in A because the focus is on the nature of the set of objects—not the nature of the individuals themselves. In D, on the other hand, one object is selected at random. Since the focus is now on one object having

---

[4] My awareness of the critical role played by conceptual boundaries in the usage of articles, and in language in general, developed to a great extent in discussions with Fred Lupke. His contribution to the theoretical basis of this paper was substantial. In places it is difficult to tell where his ideas leave off and mine begin. For further treatment of boundary phenomena see Lupke (1977). The specific applications of those ideas and the attendant problems, however, are my own.

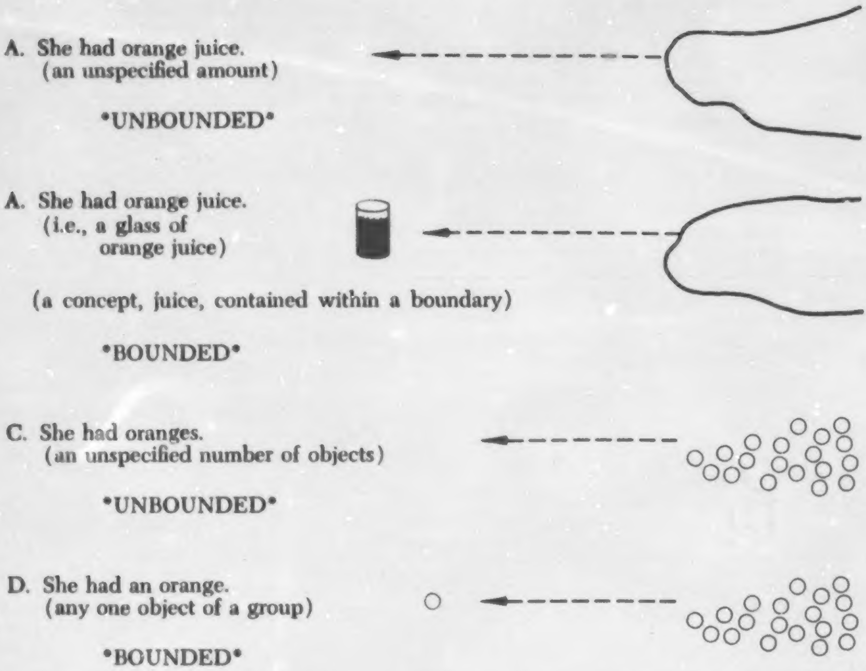

A. She had orange juice.
(an unspecified amount)
UNBOUNDED
A. She had orange juice.
(i.e., a glass of
orange juice)
(a concept, juice, contained within a boundary)
BOUNDED
C. She had oranges.
(an unspecified number of objects)
UNBOUNDED
D. She had an orange.
(any one object of a group)
BOUNDED

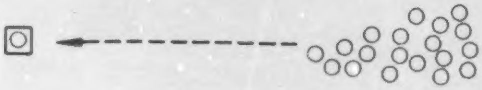
Figure Two. A Visual Representation of Bounded and Unbounded Concepts—I (the use of the indefinite article versus no article).

inherent boundaries (the skin of the orange as opposed to the amorphous form of juice) the indefinite article is used to mark the concept as bounded. D is analogous to B. The boundaries coincide with physical containers. In many cases, however, the boundaries are not physical but mental, as in (6):

(6) Boris had a vision.

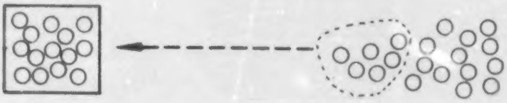
Examples E and F in Figure Three, A Visual Representation of Bounded Concepts—II (Definiteness), illustrate the second function of articles: to indicate that the speaker assumes that a given concept is sufficiently clear and well identified so that the listener can separate it from all others like it. In example D, the significant boundary is the shape of the individual orange. The boundary signaled by the indefinite article separates the physical object, an orange, from the idea of orange juice or the color orange. The boundaries in E and F function at a different level. The arrows leading from the unbounded sets on the right to the orange(s) in the box(es) on the left should be interpreted as follows: Whatever is inside the box has either been mentioned before in the conversation (or text) or somehow represents information that is

E. She had the orange.
(one object separated from all others like it)



BOUNDED

F. She had the oranges.
(a clearly identified group separated from all others like them)



BOUNDED

Figure Three. A Visual Representation of Bounded Concepts—II (Definiteness).

shared by the speaker and the listener (see below for discussion of "shared" information). The box surrounding the orange in E has nothing to do with the shape of the orange, but rather it says that what is inside is uniquely identifiable or locatable (Hawkins, 1974). The same applies to F. Suppose that F represents a subsequent reference to the unbounded concept in C, "She had oranges." The definite article tells the listener only that it is the same oranges. The exact number may still be irrelevant. F could also, of course, be a subsequent reference of an earlier mention of a specific number of oranges (i.e., "She had five oranges."). In summary, the difference between the first two functions of articles can be described using the idea of boundaries. The indefinite article, like numbers, indicates what kind of thing something is or "how many" it is. The definite article, in contrast, tells the listener "which one" of like objects a thing is. The separating function is the same in both instances, but the articles themselves do different things, operate on different levels.

Figure Four, A Visual Representation of Bounded Concepts—III (Generics) depicts the third function of articles. The problem of how (or if) to teach the generic uses of articles has been dealt with by a few linguists (e.g. Hok, 1970; Kaluza, 1971). For the most part, the strategy has been to isolate the two uses of articles, teaching them separately. Again, as was the case with the differences between the functions of definite versus indefinite articles, the concept of the conceptual boundary can be utilized in explaining the relationship between the generic and non-generic uses of articles. It is often pointed out that the generic forms are seldom used and therefore deserve little attention. The problem with that approach is that students do not then fully appreciate the fact that the slot before a noun (for instance) is meaningful—no matter what form is there. This is especially true in the contrast between the use of the definite article versus no article, as in (7) and (8):

(7) Boys are always hungry.
(8) The boys are always hungry.

G. Oranges are round.

(a boundary focussing on the existence of a group)

H. An orange is round.

(a boundary focussing on the existence of a group and a boundary focussing on one individual as a model)

I. The orange is round.

(the two boundaries in H above plus an additional boundary of separation)

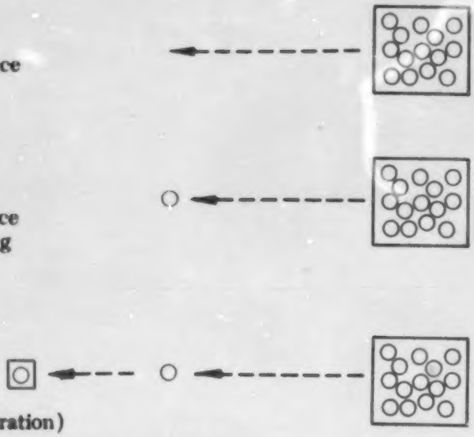


Figure 4. A Visual Representation of Bounded Concepts—III (Generics).

One must understand generics to know that those two sentences are not equivalent.

The basic difficulty in dealing with generics is that the choice of which generic to use is often very complex. As noted above, articles indicate which generic is being used. A discussion of generics in terms of boundaries helps the student understand them and also provides some guidelines as to the contexts in which each is used.

In G in Figure Four, "Oranges are round", the relevant boundary, the box, is seen to surround all oranges in existence. Here the construction imposes a boundary, establishing the existence of a set of objects. Attention is directed to the group as a whole. In H, there are two relevant boundaries: one around the group (as in G) and one around an individual orange (as in Figure Two, D). The emphasis here is on typicality, any individual as a representative of the group. The generic in I involves three boundaries: an "existence" boundary (as in G), an "any individual" boundary and an additional separating boundary (signaled by the definite article). The "existence" boundary is sufficient in some sense to separate one group from another. It creates a group that shares certain characteristics. The addition of a second separating boundary, as in I, seems to suggest that the generic in this context needs to be further separated from something. In fact there seems to be a clear tendency for generics like I, "The orange is round" to occur more frequently in passages where there is some degree of explicit comparison with a related group or set involved, e.g. comparing oranges with bananas. As an initial teaching strategy one might begin with the generalization that generics such as "An orange is round" tend to be used in descriptive (rather than comparative) discourse where the characteristics of the group are being enumerated, or where a model is being constructed. Generics of the type in G appear to be used most commonly where group attributes are being discussed and where there are no comparisons being made or models proposed; that is, in cases where the generalization is somewhat less a central element in the discourse.

A fourth function of articles is to indicate distance from the Self. The Self versus Other dimension as an explanation of certain grammatical phenomena comes from many sources including Jespersen's notion of "familiarity" and Cooper and Ross' (1975) use of distance from Self or Ego. Those two concepts will be used to account for certain cases where one has to choose between the definite article or no article—probably the most difficult set of decisions for the learner. The basic consideration is whether or not the speaker and listener are thought to be inside or outside of the boundaries of some group or institution. As will be demonstrated below, the general rule is something like: If an object(s) is conceived of as close to Self, then no article is used; if something is closer to the Other (outside a boundary), the definite article is used. For example, one of the natives would probably not say:

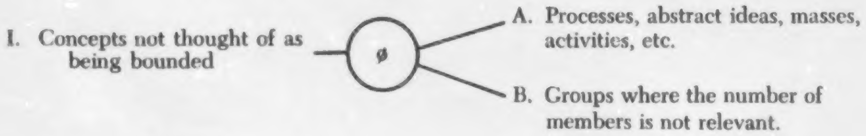
(9) The natives are restless tonight.

unless he was trying to change his identity. The definite article in (9) suggests to the listener that the speaker is not a member of that group.

Part Two: A Conceptual Framework for Article Usage

The remainder of this paper is a discussion of the categories of Figure One. When using this framework as a teaching tool, going through an expanded version of Figure One, with numerous examples, seems to be a good review for advanced students. For less advanced students only small "pieces" of the presentation in Figure One should be used—in conjunction with both inductive and deductive instruction in articles and related grammatical units.

First we will examine concepts marked as unbounded, category I in Figure One. Proper Names and Common Names are bounded. The term "name" was chosen primarily because in some cases gerunds, infinitives and adjectives—in addition to nouns—can be objectified. The choice as to whether something is to be treated as bounded is simply a matter of what the speaker wants to focus on. Proper Names and Common Names can, at least in principle, all have unbounded counterparts. Therefore, the terms "proper nouns and common nouns" are not used, to avoid the confusion of "changing" labels back and forth. In other words, the notion "noun" in this system is not somehow more "basic" than "adjective". An adjective does not function as a noun; it is either bounded or unbounded in the same way that a noun can represent either category.


I. Concepts not thought of as being bounded
ø
A. Processes, abstract ideas, masses, activities, etc.
B. Groups where the number of members is not relevant.

Category IA of unbounded concepts is made up of expressions such as:

(10) dancing, history, water, induction

all of which can be expressed as bounded (e.g., the dancing of Fred Astaire, the history of Brazil, a glass of water, the induction of John Doe). Also in this category are the unbounded concepts in the following pairs:

(11) to church/to the church, to bed/to the bed, to prison/to the prison.

Diseases and illnesses that are perceived as unbounded (not of expected short duration), for instance,

(12) polio, pneumonia, cancer, etc.

are not preceded by articles (understandably), whereas more temporary ailments on the order of

(13) a cold, a broken arm, an upset stomach, the flu

are treated as bound concepts. The reason for the definite article with "the flu" may be due to the fact that that particular illness is one that is often thought of as coming "in waves"; that is, we only deal with one current strain at a time (each strain being of relatively short duration for an individual). One rather unpleasant counter-example here is

(14) diarrhea.

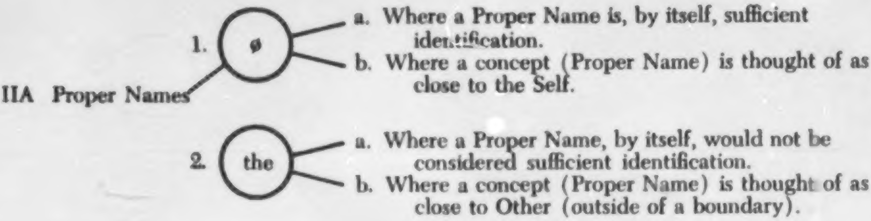
Could it be that native speakers think of it as being interminable? Notice that in the euphemisms for diarrhea, as is the case with many euphemisms, the definite article is used:

(15) the runs, the trots.

Euphemisms in relation to the grammar are "objects" in that they are separated from normal grammatical processes and semantic rules. Idioms are likewise "objects" in many cases.

Category IB, unbounded groups (where the exact number of members is irrelevant), was discussed in conjunction with Figure Two, example C, "She had oranges". Again, the key point is that the speaker has decided that the number of objects involved is either unknown or unimportant. This category also includes the use of quantifiers with groups of objects, e.g., many, a few, a lot of. Note that category IB only applies to the first mention of a concept. When what was initially expressed as unbounded is mentioned again it will be marked as bounded, as in Figure Two, example F, "She had the oranges".

Next we move to the other basic category in the framework, bounded concepts. The main reason for making the first distinction in category II the opposition between proper and common names was that proper names generally have a uniqueness by virtue of the name itself (e.g. Detroit) and are considered bounded. That uniqueness carries over even to proper names that are preceded by the definite article. A student can say, "I live in United States" and be understood because of the context. Proper Names constitute a large number of "exceptions" in most grammars. The idea of boundaries helps to explain many of those "irregularities." First notice that A1 is the converse of A2.



a. Where a Proper Name is, by itself, sufficient identification.
b. Where a concept (Proper Name) is thought of as close to the Self.

a. Where a Proper Name, by itself, would not be considered sufficient identification.
b. Where a concept (Proper Name) is thought of as close to Other (outside of a boundary).

Category A.1.a., proper names with sufficient identification such that they do not ordinarily need an additional boundary (signaled by a definite article), would be names such as:

(16) George Washington, Ringo Starr.

However, a proper name of that type could still require an article if, for instance, there were more than one Ringo Starr in the domain of the discourse. In that case A2a would apply (i.e., The Ringo Starr who sang with the Beatles), since without the extra boundary that person could not be identified. Other examples where A2a is relevant include:

(17) The Rocky Mountains, The Venezuelan Government, The Detroit River.

Without the article, none of the concepts in (17) are bounded. As was the case with "United States" above, the chance of a student being misunderstood is slight. However, he should understand the principle involved. The second example in (17) is interesting. Without the article it would refer to an unbounded process, a way of doing things—not to the official government of the country. "The Detroit River" is also revealing. What is the best way to explain why we must have the definite article in that expression? It can be handled rather neatly using the concept of boundaries, along with the model of a computer trying to decipher that noun phrase. If the computer were to come to a word like "Detroit", it would automatically tag it as a bounded concept, a proper name—not a modifier. Without the introductory article, the computer would have to go back and reinterpret "Detroit" when the word "river" appeared. The presence of the definite article, however, would signal that what follows *needs* an extra boundary, thereby anticipating the word "river" or some other object.

Next consider the difference between A1b and A2b, close to the Self versus close to Other. A1b is concerned, primarily, with cases where a name is normally preceded by the definite article but which, in everyday usage, are often said without the article. These are a particular cause of confusion for students. Examples (18) through (21) illustrate the problem:

(18) I work at The University of Michigan
*(19) I work at University of Michigan.[3]
*(20) I work at the U. of M.
*(21) I work at U. of M.

The rule as to which of those four forms one uses in conversation involves a consideration of whether the speaker and/or listener are in some way thought to be familiar with or within the boundaries of that institution. For instance, if one were in Ann Arbor, talking to someone else from that same city, he would probably use either (19), (20) or (21)—as opposed to (18). There are two principles at work in those examples. The first principle, discussed in Platt (1974) and elsewhere, is that you only use abbreviations with other people who are familiar with the institution and where there would be no resultant ambiguity. Second, the definite article signals that what follows may be separated, or distant, from something else. To use (18) with a sales clerk, for example, at a store in Ann Arbor might very possibly be perceived as an attempt to sound important, putting more social distance than necessary between the customer and the clerk. It almost says, "I need to give you the full title since I assume you may not be sufficiently familiar with the shorter form(s) to recognize what I am referring to." To use no article is to share a level of informality, to imply that—a least on this matter—we are both within some boundary (social or pertaining to mutually shared information). If one were distant from Ann Arbor, conversing with someone relatively unacquainted with The University of Michigan, he would probably use (18). Examples (22) and (23) are closely related to (18) through (21).

(22) Democrats are in control of both houses.
(23) The Democrats are in control of both houses.

When would a Democrat say (23)? As was the case in (9), "The natives are restless", normally, if you are a member of a group—within the boundary of the group—you do not use the article unless you are trying to sound "objective". In (23) the speaker wants the listener to think that he is in some way "outside" of the group. In (22) the speaker might be either a member of the group or he might simply be signaling that it is unimportant to him whether he is considered "one of them".

Another example of the distancing effect is the case of

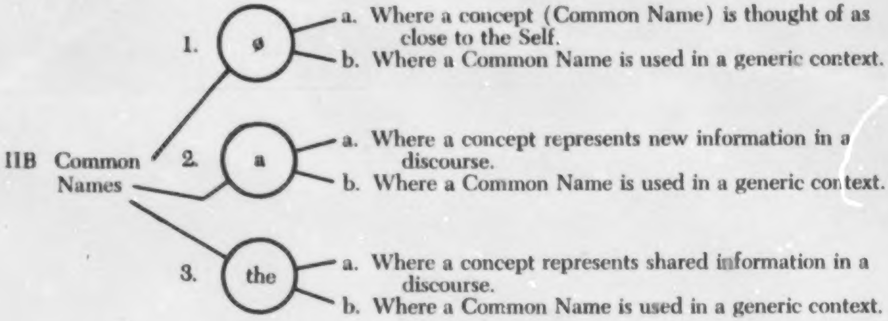
(24) The Fonz

on the television program "Happy Days". The character's name shifts in certain situations from "Fonz" (short for "Fonzerella") or "Fonzie" (the intimate form) to (24) as his status changes from "one of the gang" to superstar. In fact, the character, himself, seems to have been the one to establish the title, The Fonz, by referring to himself in that manner whenever he did something spectacular. The extra boundary serves to set him off from ordinary mortals.

---

[3] The asterisk is used here to refer to examples that are considered ungrammatical or of questionable grammaticality to some speakers.

The other major category of bounded concepts, IIB, is Common Names.


IIB Common Names
1.
ø
a. Where a concept (Common Name) is thought of as close to the Self.
b. Where a Common Name is used in a generic context.
2.
a
a. Where a concept represents new information in a discourse.
b. Where a Common Name is used in a generic context.
3.
the
a. Where a concept represents shared information in a discourse.
b. Where a Common Name is used in a generic context.

Observe that in the Common Names category there are basically three distinctions or oppositions: generic/non-generic, new/shared information and close to Self/Other.

Category IIBla is the common name counterpart of IIAlb (proper names with no article preceding them—seen as close to the Self.) Compare (25) with (26):

(25) Teachers can not agree on which method is the best.
(26) The teachers can not agree on which method is the best.

A teacher might use (26) in a special context; for instance, if he were serving as a representative for a group of teachers. By signaling that he considered himself to be outside of the group for a specific purpose, he would, in effect, be demanding to be treated differently. The power of shifting boundaries is very clear when a spokesperson for a crowd suddenly changes his point of reference in his speech from "we" to "the people".

Category 1b, generic contexts where the simple plural form is used, has been touched upon in example G, Figure Four, "Oranges are round." Although the plural noun is used, referring to things that could be counted, the focus is more on membership in a group than on the nature of the individual members themselves or the characteristics of this group relative to some other group.

The basic consideration involved between the use of the definite and indefinite articles in categories IIB2 versus IIB3 is whether uniqueness is seen by the speaker as important or not. That is crucial in situations such as (27) and (28), assuming that both sentences could serve to introduce the same object (the boy) into the conversation for the first time, and that the listener had no previous knowledge of the boy:

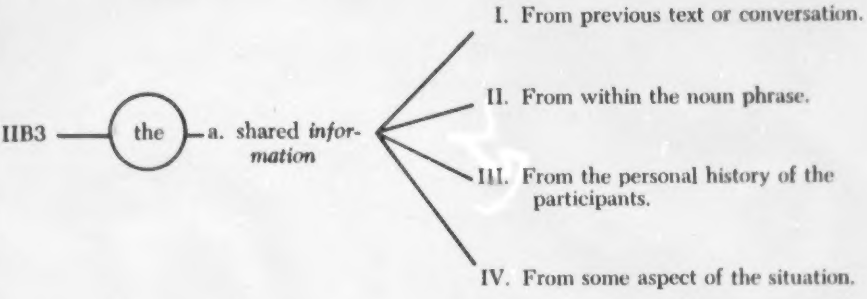
(27) A boy who lives down the street is sick.
(28) The boy who lives down the street is sick.

In both cases the subject is followed by a relative clause, but for some reason the speaker felt it was necessary to designate the subject in (28) as being unique. Some of the conditions for determining uniqueness will be listed in the discussion of B3a.

B2a represents the "first mention" or "new information" function of the indefinite article. If uniqueness is unimportant and it is the first reference to a common name, a concept will generally be preceded by an indefinite article as in (27).

B2b is what has been earlier referred to as the generic use of the indefinite article (example H in Figure Four, "An orange is round"). The reason for putting the generic function here, as a category of the common name function, is that the speaker, by selecting this form of the generic, has chosen a bounded object—a common name—to be an example of a bounded group. The choice of which member, the uniqueness decision, is not relevant.

Category B3, the use of the definite article with common names, is the most complex. Like B2, it is divided into two subcategories, one related to the status of information (shared information), the other related to generics. B3a represents the use of the definite article in cases where the speaker and the listener share enough information about a common name to adequately identify it in that particular discourse. There are at least four sources of that information:


I. From previous text or conversation.
II. From within the noun phrase.
IIB3
the
a. shared *information*
III. From the personal history of the participants.
IV. From some aspect of the situation.

I. *From previous text or conversation.* That is what is commonly referred to as the "previous mention" function, cases where the object has been *explicitly* introduced prior to the "time" when it is again referred to. Whitman (1974) reported that this function accounts for only about fifteen percent of the uses of the definite article. It is, nonetheless, generally taught first.

II. *From within the noun phrase.* In this case the information needed to uniquely identify an object comes from a construction such as a relative clause, a prepositional phrase, a post-posed adjective:

(29) the cat that ate the salami, the son of John Doe, the president elect.

As discussed earlier, just because the information that could potentially make an object unique is in the sentence (or noun phrase) does not necessarily mean that the speaker will treat it as unique, as in examples (27 and (28). Sometimes the logic of the constructions dictates uniqueness:

(30) the best, the least memorable, the only child at the party.

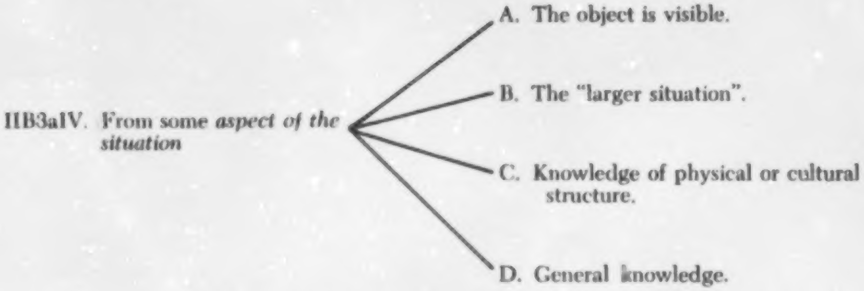
There are many aphoristic expressions (and idioms, as discussed earlier) that require the definite article:

(31) the more the merrier, the early bird gets the worm.

III. *From the personal history of the participants.* This category refers to private, shared knowledge that could not be correctly interpreted by "outsiders" on linguistic or cultural grounds alone:

(32) How is the leg coming?

IV. *From the situation.* Hawkins (1974) breaks down the situation into context types. For this analysis we will use only four of his categories, collapsing some functions that Hawkins, for theoretical reasons, deals with separately.


IIB3aIV. From some *aspect of the situation*
A. The object is visible.
B. The "larger situation".
C. Knowledge of physical or cultural structure.
D. General knowledge.

Category A is where the object is visible to both the speaker and the listener:

(33) Please pass the nutcracker, Sweet.

As an example of category B, Hawkins cites a hypothetical situation where two people approach a fence and one says:

(34) "Watch out for the dog."

even though the dog is not in sight. Another example would be:

(35) I went to the firestation this morning.

In both instances, the situation dictates there was (or is going to be) one, unique object involved. The situation plus the definite article gives the listener all the information he needs. Category C reflects knowledge of physical or cultural structure. For instance, if an automobile has been introduced then almost every part of the automobile can be referred to as unique (bounded):

(36) They ripped off the horn, the muffler, the tires and the radiator.

The same goes for a wedding:

(37) The bride, the groom, the ringbearer and the ushers left after the dinner.

Category D, general knowledge; concerns common names of things that are thought of as being inherently unique:

(38) the sun, the moon, the air, etc.

The last category of common names, IIB3b, deals with generics that make use of the definite article. Example I in Figure Four, "The orange is round", was discussed earlier. It may be helpful for the student to think of the definite-generic as imposing another boundary on an indefinite-generic, allowing the possibility of a stronger contrast. Observe the difference between

(39) A rich person eventually gets a facelift.

and

(40) The rich person eventually gets a facelift.

(40) seems to be a potentially better form to use in a discourse dealing with comparison, e.g., comparing the rich with the poor. Also, consider the relationship between (41) and (42):

(41) The cat has whiskers. (in the generic sense)
(42) The cats have whiskers. (also in the generic sense)

(41) can be understood as referring to housecats, but (42)—not a generic in the usual sense—seems to refer only to the set of species of cats, e.g., lions, tigers, housecats. (42) almost invites a comparison with some bounded group of equal biological rank. It is almost as if (42) has one more boundary imposed on it than does (41).

Some Conclusions and Suggestions

This brief overview, or "map" of some aspects of the use of articles is based on several assumptions as to how grammar–and articles in particular—should be taught. In the framework presented in Figure One, the visual and spatial features of the diagram are assumed to be beneficial to the learner in helping him literally *see* the relationships involved. It is "assumed" because this approach has not been tried without the accompanying graphic displays. The whole idea of boundaries would be impossible to explain without some kind of visual support. It would be similar to teaching without a chalkboard.

There is some anecdotal evidence of the reality of the framework for students. More than one student, when trying to explain why an article belonged in a particular environment has unconsciously pointed in the air to the mental image of the framework he had in his mind—a week or more after the framework had been presented in class.

The Self versus Other opposition, as presented in conjunction with the idea of boundaries, has generated considerable interest among students. It is a case where the grammar seems to be reflecting something of the way English speakers categorize the world. To be able to choose between using no article or the definite article often requires a thorough understanding of the native speaker's concept of Self. Many cultures are much less prone to putting boundaries on concepts. To isolate or fence in an idea, that is, to treat it as an object, may be to think about it very differently (Whorf, 1951). For many students of English as a second language that process is quite alien to their own way of categorizing experience. Once understood, however, it may help to explain some otherwise irrational linguistic and social behaivor on the part of native speakers.

A key question in teaching articles has to be: Exactly how important are articles relative to other structures? In terms of the student's ability just to communicate basic information, articles are probably less important. How much time should be allotted to making the student be able to use articles accurately? Many studies have shown that articles are acquired not in isolation, but rather in conjunction with related structures. That would imply that one important teaching strategy is stressing the strong interrelatedness of articles with other garmmatical units. In fact, one of the purposes of this type of framework and the research on boundaries is to show how various processes are related both functionally and formally.

Teaching articles to students of various levels of English proficiency makes one aware of an apparent paradox. Advanced students are capable of benefitting from conscious attention to the pragmatics or situational aspects of articles (and other grammatical processes) in ways that are well beyond the beginning student. Advanced students can be asked to verbally create contexts in which almost any isolated sentence might be appropriate. If handled correctly, that sort of exercise (context-creating exercises) can be a rich source of grammar, conversation and vocabulary. The beginning student, on the other hand, can not very well deal with sentences out of context when working on articles, etc. The contexts must be provided. Explicit, communicative situations seem to provide the best teaching tools at that level. One general model of the learner, accepted by many teachers and theorists (if only tacitly), is that the learner should move from mechanical activities toward more natural, communicative activities. Accordingly, the input to the learner should change from isolated sentences to longer discourse. Ironically, when it comes to teaching articles, at least, the most effective focus for teacher input to the learner may instead progress from discourse to isolated sentences.

REFERENCES

Barrie, William. 1971. What's in a determiner? *Les Langues Modernes* 65, 4, 69–80.

Christophersen, Paul. 1939. The articles: a study of their theory and use in English. Copenhagen, Munsksgaard, Oxford University Press.

Cooper, W. and J. Ross. 1975. Word Order. *Papers from the parasession on functionalism*. Chicago Linguistic Society. Pp. 63–111.

Grannis, O. 1972. The definite article conspiracy in English. *Language Learning* 22, 2, 275–89.

Hawkins, J. 1974. Definiteness and indefiniteness. Ph.D. Thesis, University of Cambridge, England.

———. 1976. On explaining some ungrammatical sequences of article + modifier in English. *Papers from the thirteenth annual regional meeting of the Chicago Linguistic Society. Chicago.* Pp. 287–301.

Hok, Ruth. 1970. The concept of "general-specific" and its application to the/a and some/any. *TESOL Quarterly 4*, 3, 231–39.

Jespersen, Otto. 1949. *A modern English grammar*. Copenhagen, Munksgaard.

Kaluza, H. 1971. Two functions of English articles. *Workpapers, Teaching English as a Second Language* 5, 69–74.

Leech, G. and J. Svartvik. 1975. *A communicative grammar of English*. London, Clowes & Sons/Longman Group Ltd.

Platt, John. 1974. Alphabet soups or a mess of pottage? *Foundations of Language 11*, 2, 295–97.

Tyma, Steve. 1977. Time, aspect and modality in English. Unpublished paper. University of Michigan.

Whitman, Randal. 1974. Teaching the article in English. *TESOL Quarterly* 8, 3, 253–62.

Yotsukura, Sayo. 1970. *The articles in English: a structural analysis of usage*. Mouton, The Hague.

Listening Skills in the Elementary ESL Class

Ana Uhl Chamot

University of Texas at Austin

Competence in listening is essential for children learning English as a second language if they are to participate fully in English speaking situations. The level of competence needed is greater than simple recall of material.

A Listening Model incorporating a hierarchical sequence of listening skills describe six cognitive levels involved in developing listening competence: discrimination, recognition, recall, analysis, inference and evaluation. The Listening Model can be used to select and classify different ESL listening activities so that a sequential program can be planned for elementary school use.

A set of criteria for designing and/or adapting ESL listening activities for the elementary school is described. These criteria are based on relevant research and methodology drawn from two primary sources: secondary level foreign language listening comprehension approaches, and elementary language arts listening programs for native speakers of English.

The goal is to provide the classroom teacher with guidelines and a framework that can be used to design an effective listening component for the ESL program.

Rationale for Teaching Listening

The skill of listening is often neglected in the classroom, yet students spend a large proportion of their time in school listening (Logan, Logan and Paterson, 1972, pp. 44-45). This is particularly true at the elementary school level (Wilt, 1975, pp. 626-636). Younger children have to learn a great deal through their ears because their reading skills are still at the beginning stages (Donoghue, 1971, p. 145). The child who is acquiring English as a second language needs to develop listening comprehension skills not only to understand what is going on in the class, but also to participate in English speaking situations outside the school.

Most English as a foreign language textbooks have some listening comprehension exercises, but frequently the listening practice provided is limited both in quantity and in variety. Specific instructional materials for improving aural comprehension are available for secondary and adult learners (Morley, 1975), but not for the younger student whose needs, interests and abilities are different. The elementary ESL teacher often has to devise supplementary listening activities if the listening skills of children who are learning English as a second language are to be fully developed.

We all know that there are different kinds of listening. The kind of aural attention we expend when listening to background music is quite different from the kind of attention given to an interesting lecture or to the filling station attendant's directions when we are lost in strange city (Donoghue, 1971, pp. 145-149). For purposes of language learning, we concentrate as teachers on the skills of attentive listening, with the immediate goal that of comprehension (Chastain, 1976, p. 278). Once the second language learner can comprehend the meaning of an oral message, we would then like to develop the same listening skills needed by the native speaker, that is, listening for information and critical listening, which implies evaluation of the oral message (Weaver and Rutherford, 1974, pp. 1146-1150; Tutolo, 1975, p. 1108).

In the literature on developing listening skills in a second or foreign language, a hierarchy of skills is generally postulated with suggested types of listening activities for each level (Chastain, 1976, pp. 293-301; Rivers, 1975, pp. 93-104; Valette and Disick, 1972, pp. 141-142). A typical hierarchy describes the lowest level listening skill as the ability to discriminate the sounds of the new language. The next higher level concerns perception of the message contained in a word or group of words. The third level involves auditory memory, the ability to remember meaningful sequences of speech. The fourth level is generally described as that of comprehension. Some authors feel that initial comprehension must necessarily be through the mediation of the first language, and that comprehension directly in the second language and long term retention is characteristic of the fifth level in the listening skill hierarchy (Chastain, 1976, pp. 291-292). In teaching a foreign language the goal for many teachers is this fifth level, and when students can successfully recall the factual information they have listened to and respond to it in the target language, this goal has been attained.

In the literature on developing the listening skills of elementary school native speakers of English, similar kinds of skill hierarchies have been proposed (see Logan, Logan and Paterson, 1972, pp. 55-56; Ramanauskas, 1974, pp. 1142-1145; Weaver and Rutherford, 1974, pp. 1149-1150). A good example is one developed by Kellogg (1971, pp. 127-128). In it the lowest level is acuity, the ability to hear normally. The second level is discrimination and the third is comprehension. Kellogg's main interest is at the comprehensive level, which he further divides into four hierarchical levels. The lowest of these requires the listener to recall the data heard, and appears to be similar to the third level of auditory memory described in the foreign language listening skill hierarchy. The next three levels of Kellogg's model require increasingly complex cognitive operation of analysis, synthesis and applications. Certain types of questions are provided which serve to illustrate the comprehension level at which the listener has received the oral message. A similar analysis of the comprehension level of listening is provided by Tutolo (1977, p. 263), who divides it into components of literal comprehension, interpretations, and critical listening or evaluation.

An ESL Listening Model

Comprehension beyond the recall level is often not a realistic goal for the learner of English as a foreign language, particularly if the environment is not an English speaking one. However, the child learning English as a second rather than as a foreign language will eventually be using English as a language of instruction, and therefore needs to develop listening skills beyond the recall level in order to compete successfully in an English language environment. A synthesis of the two types of listening skill hierarchies described above can provide for this child a sequence of listening skills ranging from initial sound discrimination, through recall of information, and on to the higher levels of critical listening.

Table 1 presents a proposed hierarchy of listening skills that provides a sequential plan for teaching listening in elementary school ESL programs.

TABLE 1

ESL Listening Model.

Level	Cognitive Process	Student Behavior
6	Evaluation	a. affective interpretation b. transfer to real life situations
5	Inference	a. drawing conclusions b. predicting outcome c. understanding new words from context clues
4	Analysis	a. identifying main idea b. selecting principal points
3	Recall	a. remembering details b. remembering sequence
2	Recognition	a. identifying word meanings b. identifying recombination sentences
1	Discrimination	a. perceiving sound contrasts b. perceiving word contrasts c. perceiving contrasts in sentence types

The first level skill involves discrimination of sounds both in isolation and in words, and discrimination of sentence patterns. At the second level the listener must recognize as meaningful both individual words and sentences containing familiar elements. The third level requires detailed and sequential recall of material heard. At the fourth level the listener begins to think analytically about what has been heard and must select the principal points and the main idea. At the fifth level skills of inference are used to listen between the lines in order to understand implied meanings. By the sixth level the listener is able to evaluate what is heard and to interpret it in terms of personal experience.

The student behaviors described for each level represent types of objectives for listening activities involving the six levels of cognitive processes. The behaviors specified are representative only, as additional behaviors can be elicited at each of the levels.

The Listening Model can be used to select and classify different kinds of listening activities, thus establishing a teaching sequence for them. Each skill level builds on the preceding ones, so that the acquisition of listening skills is cumulative.

Criteria for Listening Activities

In selecting or developing listening activities certain criteria should be established in order to ensure that a listening activity is a learning rather than a merely recreational experience. Various criteria have been proposed both by foreign language methodologists (Chastain, 1976, p. 286; Rivers, 1975, pp. 60, 76-78, 91-92) and elementary school language arts specialists (Hollingsworth, 1972, pp. 267-268; Logan, Logan and Paterson, 1972, pp. 55-56; Donoghue, 1976, pp. 143-144).

The first criterion in selecting an educationally sound listening activity can be accomplished by the identification of its learning objective (Hollingsworth, 1972, p. 267). A description of what the child must do in the activity can be matched to the cognitive process it reflects, as identified on the Listening Model. If the child will be discriminating same and different sounds, then the learning objective will be found at Level 1 of the Listening Model. If the child will have to recall the plot of a story, then the learning objective will be found at Level 3. If there is no apparent match between the proposed behavior and an identifiable cognitive process, then the teacher should question the specific instructional value of the listening activity, as many so-called listening activities resemble party games more than planned learning experiences.

The second criterion involves selection of material at the appropriate interest level. Even though a fifth grader and a first grader are at the beginning stage of learning English and need to develop the ability to discriminate contrasting sounds, they will probably not want to engage in identical listening activities. The additional maturity of the older child as well as the fact that he is literate in his first language will make it possible for him to engage in a more sophisticated listening activity than is possible for a preliterate younger child. Listening selections and the visuals that accompany them should be meaningful to the children who use them. Younger children may be interested in animals and storybook characters, whereas older children would find comic book and television heroes more appealing.

A third criterion to consider is the length of the listening activity. Young children have short attention spans and it is difficult for them to sit quietly for extended lengths of time. Active listening should not last longer than a few minutes, and responses to listening should be as varied and as active as possible.

Short listening activities can be repeated with variations frequently in order to assure adequate practice, but children can become restless if they are expected to listen intently for more than a few minutes at a time.

A fourth criterion concerns the use of other language skills in the listening activity. Young children who are just beginning to read and write should engage in completely oral activities in developing their listening skills, but older children who are reading and writing with some ease may profitably combine these skills with a listening activity (Rivers, 1975, pp. 76-77). However, the teacher should evaluate children on the listening aspect of the activity, not on speaking, reading or writing ability. Because it is often difficult to separate the skills in determining whether a difficulty is caused by inadequate ability in listening or in one of the other skills in a given exercise, it is often more satisfactory to use listening exercises in which the responses involve physical movement, identification of pictures, drawing or making a paper with numbers or symbols.

The criteria described so far are as applicable to listening activities for native speakers as they are for learners of English as a second language. Additional criteria for the ESL class would include the desirability of redundancy in the listening message (Rivers, 1972, p. 89). In normal speech a great deal of redundancy is provided both by the structure of the language and by the paralinguistic behavior of the speaker. All of this helps the listener to understand: if one part of the message is missed, it is generally repeated in some way. Foreign language listening exercises often lack this feature of redundancy, making comprehension more difficult than it has to be. Redundancy should be consciously included in all listening activities. Because non-verbal behavior greatly assists the communication process, the teacher should transmit difficult messages personally rather than through the tape recorder and should use the gestures, body movements, pauses and intonations that normally accompany oral messages. Listening activities on tape probably need additional verbal redundancy, and should not be as difficult as the listening activities led by the teacher personally.

A sixth criterion for listening activities has to do with the pace at which the listening selection is read or recorded. The ideal pace has frequently been described as that normal to the native speaker, yet teachers have objected to this recommendation, feeling that learners of a new language cannot easily understand the natural pace and speed of a native speaker. However, when the normal pace is slowed too much, the result is an artificial sounding segment of language without natural intonation contours. The solution lies in lengthening the pauses between phrases and sentences. These pauses allow the listener to process what has just been heard (Rivers, 1972, p. 96), but since they occur at phrase and sentence boundaries, the natural intonation pattern of the language is not altered.

The seventh criterion can be phrased as a question: Does the pupil know what to listen for? By telling children what they should focus their attention on, increased listening skill will be more easily attained. Thus, if the objective is to listen for details, the teacher might tell the children before reading the listening selection, that they should try to remember each thing that happens in the story. If the objective is to guess meaning of new words from context, the teacher might write the new words on the board before the children hear the listening selection. In this way the teacher provides pupils with advance organizers so that they will be mentally set to employ appropriate cognitive strategies in the listening activity that is to follow.

Finally, the use of appropriate visual aids is advisable for most listening activities. Children can attend to the listening selection more easily if they have something relevant to look at while they are listening. They can then relate what they perceive through the visual channel to what they are perceiving through the aural channel, and the learning experience becomes more meaningful. Needless to say, visual aids should not distract from the oral message, but should provide a focus for the listener's attention.

Summary

Children learning English as their second language need to develop listening skills in order to function successfully in English speaking environments. The type of listening skills needed goes beyond the recall level frequently set as a goal for the foreign learner.

A synthesis of listening skill hierarchies of foreign language and elementary language arts teaching theories is proposed. This ESL Listening Model provides a six level sequence of cognitive processes inherent in listening, from the beginning level of discrimination to an advanced level of evaluation. The Listening Model describes the cognitive process for each level and lists representative student behaviors that can be elicited. Teachers can use it to select and devise listening activities with specific learning objectives for each level.

In addition, a set of criteria for selecting, preparing and presenting listening activities is proposed. The purpose of these criteria is to aid elementary teachers in planning and implementing an educationally and linguistically sound listening component for their ESL program.

REFERENCES

Chastain, Kenneth. 1976. *Developing second-language skills: theory to practice.* Second Edition. Chicago, Rand McNally.

Donoghue, Mildred R. Listening. 1971. *The child and the English language arts.* Dubuque, Iowa, Wm. C. Brown.

Hollingsworth, Paul M. 1972. Teaching listening in the elementary school. In H. Funk and D. Triplett (eds.). *Language arts in the elementary school: Readings,* Philadelphia, J. B. Lippincott.

Kellogg, Ralph E. 1971. Listening. In P. Lamb (ed.). *Guiding children's language.* Second Edition. Dubuque, Iowa, Wm. C. Brown.

Logan, Lillian M., V. G. Logan, and L. Patterson. 1972. Perceptive listening. In

Creative communication: teaching the language arts, Toronto, McGraw-Hill Ryerson.

Morley, Joan. 1975. *Improving aural comprehension*. Ann Arbor, University of Michigan Press.

Ramanauskas, Sigita. 1974. The content, programs and teaching dimensions of aural language. *Elementary English 51*, 8.

Rivers, Wilga M. 1972. Listening comprehension. In Kenneth Croft (ed.), *Readings on English as a Second Language,* Cambridge, Mass.

Rivers, Wilga M. 1975. *A practical guide to the teaching of french*. New York, Oxford University Press.

Tutolo, Daniel J. 1975. Teaching critical listening. *Language Arts* 52, 8.

Tutolo, Daniel J. 1977. A cognitive approach to teaching listening. *Language Arts 54*, 3.

Valette, Rebecca M., and R. S. Disick. 1972. *Modern language performance objectives and individualization: a handbook*. New York, Harcourt Brace Jovanovich.

Weaver, Susan W. and W. L. Rutherford. 1974. A hierarchy of listening skills. *Elementary English 51*, 8.

Wilt, Miriam E. 1950. A study of teacher awareness in listening as a factor in elementary education. Summary in *Journal of Educational Research 43*.

Tagmemic Concepts in the ESL Speaking/Pronunciation Class

William Crawford and Beverly Fried

English Language Institute
University of Michigan

This paper deals with application of tagmemic concepts to the Teaching of English as a Second Language. Specific application of the concepts of (1) *impact*, (2) *perspective*, (3) *hierarchy*, and (4) *variation*, to high-level speaking/pronunciation classes will be discussed.

Treating meaning as *impact* of a language unit on the listener (with form as the physical content of the unit) students saw the result (*impact*) of their utterance (or other unit of discourse) on the listener.

Applying the notion of *perspective* to speaking skills, students were led to examine English phonology from three viewpoints: the *static*, *dynamic* and *relational* perspectives. The *static* (or *particle*) perspective allowed us to look at the sound system of English as being made up of individual particles, elements, or items. At this level of analysis the focus was on the contrastive-identificational features that distinguish individual sound segments, their variation and physical manifestation, and their distribution in class, sequence and system. The *dynamic* (or *wave*) perspective focused upon "the stream of speech" as a single dynamic, moving entity. Attention was drawn to variation due to speed of utterance, placement in a stress group (or other phonological unit), and spread of phonetic features (via shared element). The *relational* (or *field*) perspective allowed students to transcend the traditional boundaries of speaking skills and to consider such notions as temporal set, physical space, social space, mental and presuppositional space, and speaker-addressee-listener orientation.

The presentation of the specific application of tagmemics includes the development of: self-monitoring skills, teacher-student error analysis of student speeches, purposive pronunciation through the use of television games (adapted for classroom use), games stressing social interaction, and material presentation (paradigm and matrix display).

Students in our advanced level speaking classes showed continued inability to correct their own pronunciation errors, and, at the same time, were unaware of the consequences or impact of their verbal behavior (on listeners). Materials based on either traditional or generative approaches to language were not successful in providing students with the necessary tools for self-monitoring and self-correcting, which would enable them to perform meaningfully in social interactions with native speakers.

There is a current linguistic approach which provides us with a theoretical justification for entering into a discussion of language on the level of social interaction, rather than a less meaningful "word", "phrase", or "utterance" level. This theory is *TAGMEMICS*, developed by Kenneth L. Pike. With the possibility of treating *meaning* as the *impact* of a language unit on the hearer's understanding, perception, or behavior together with the *form* as the physical manifesting *content* of that unit, the students can see the result, or impact, of their utterances (or other unit of discourse) on the listener. "Intelligibility" is of primary concern here and we support J. C. Catford's conclusion (1967, p. 149) that "*intelligibility losses* are due to defective *selection* or *execution* on the part of a speaker, or to defective *identification* or *interpretation* on the part of a hearer, or to a combination of these factors." Catford stresses that (1967, p. 150) "in assessing the relative 'importance for intelligibility' of different linguistic forms, it [is] . . . most important, to direct our attention chiefly to the *speaker's* role, assuming the hearer to have a *medium threshold of intelligibility*; and to investigate the *functional value* of forms and their *frequency in crucial contexts*." Tagmemics places perhaps even greater emphasis on the role of intelligibility in social interaction. We feel that our students begin to understand the implications of their linguistic behavior when they see the value of correct pronunciation in relationship to the entire communicative act.

Some of the basic concepts of tagmemic theory which we will outline include: the *particle*, *wave*, and *field* perspectives; variation; change via shared element; purpose; and the form-meaning composite. We will examine the applications of these principles to the ESL speaking/pronunciation class in terms of: 1) presentation of materials; 2) error analysis; and 3) the use of games via impact.

Implications of Tagmemic Theory

The scope of tagmemics is not limited to linguistic behavior alone. The tenets of tagmemic theory are based upon principles that are universal to human behavior. In Kenneth and Evelyn Pike's *Grammatical Analysis*[1] the principles of tagmemic theory that constitute 'universals in the structure of human behavior' are presented in outline form (Pike and Pike, 1977:1:1):

1.1 UNIT
  1.1a contrastive-identificational
  1.1b variation and physical manifestation
  1.1c distribution
    1.1c1 in class
    1.1c2 in sequence
    1.1c3 in system

1.2 HIERARCHY
  1.2a referential
  1.2b phonological
  1.2c grammatical

---

[1] Page numbers cited in this article refer to the *prepublication copy* of *Grammatical Analysis*. There may be discrepancies in page numbering when referring to the actual text (to appear in 1977).

1.3 CONTEXT

- 1.3a form-meaning composite
- 1.3b change via shared element
- 1.3c universe of discourse

1.4 PERSPECTIVE

- 1.4a static (or particle)
- 1.4b dynamic (or wave)
- 1.4c relational (or field)

In tagmemics, "the *observer* becomes an element of the theory; no fact is treated without reference to him, so that the theory does not discuss 'the thing in itself'. . . . Since the observer can change his viewpoint, modifying the manner in which he looks at the data impinging on him, so also the theory itself must have *multiple starting points* once the observer is part of the 'given'" (Pike and Pike, 1977:1.6).

This notion of *perspective* is central to all tagmemic analysis. Applying this principle to the teaching of speaking skills allows us to examine phonology from a multi-perspective viewpoint. The *static* (or *particle*) perspective allows us to look at the sound system of English as if it were made up of individual particles, elements, or items. At this level of analysis the teacher can focus upon the individual sound segment problems that hinder a student's accuracy or fluency. The individual phoneme, then, is the center of focus from the static perspective.

Individual sound units are identified by contrastive-identificational features. A /p/ is distinct from a /b/, for example, due to its contrastive voicing and different from an /f/ by its occlusiveness and is a phoneme by virtue of filling a marginal slot of a syllable. A speaker of Arabic must learn this distinction when learning English if he is to communicate effectively.

When viewing language from the static viewpoint, we must consider how the individual sound units are distributed in class, sequence, and system. English has three nasal sounds: the bilabial stop [m], the apico-alveolar stop [n], and the velar stop [ŋ]. However, the distribution of these sounds is not the same. In English, the velar nasal [ŋ] does not occur in initial position. This fact of distribution can be very important for some students learning English pronunciation. Many non-Indo-European languages, for example, have no such phonological constraint.

Contrastive-identificational features (distinctive voicing in /p/ vs. /b/), variation (aspirated vs. non-aspirated allophones of a single phoneme), and distribution (the nonoccurrence of initial velar nasals in English) are all aspects of analysis from the static viewpoint.

The *dynamic* (*or wave*) perspective looks at linguistic phenomena (Pike and Pike, 1977:1.6) as a "series of 'discrete' events and treats the *whole* as a single dynamic moving entity. . . . Any single unit can be viewed dynamically, as having [a] beginning (*initial margin*), [a] middle (*nucleus*), [and an] end (*final margin*). In such an instance, the unit is viewed as a *wave*."

Both the phrase *'as you'* and the sentence *'Ed had edited it'* are best viewed from a dynamic viewpoint. An analysis of these examples from the static viewpoint would face difficulties in explaining change via shared element or in assigning meaning to specific phonemes. However, from the dynamic perspective we are able to treat the unit as a whole and perceive its meaning as a single unit. Pike comments on this specific issue in his new television series *Pike on Language* (produced by the University of Michigan Television Center, 1977). In the third film of the series, "Waves of Change", Pike (1977) asks the viewer:

> Suppose I ask you what this means, and pronounce it as: [ǰɪnǰoiɪt]. Perhaps you'd catch on that [ǰɪnǰoiɪt] slow is *'Did you enjoy it?'* But the *'Did you'* goes to [jɪn]. Or suppose I say [skuit], what does that mean? Well, it means: *'Let us go eat'*. But this one which is even prettier which is this: [ɛdədɛ́dətədɪt]. It's quite normal for my American English. Probably for yours, too. But, what is it? This says: *'Ed had edited it'*. But in spite of change there is some type of identity which is retained.

The *relational* (or *field*) perspective allows us to "eliminate from the center of [our] attention the form or content or extension of the units as such and focus instead on the relationships between them. The units in this case, contract to a point in a *network* of relationships" (Pike and Pike, 1977:1.7). From this perspective we are able to transcend the traditional boundaries of "speaking skills" and include for consideration such notions as temporal set, physical space, social space, mental and presuppositional space, emotional and attitudinal space, and speaker-addressee-listener orientation. Any treatment of pronunciation that excludes such considerations as these neglects the actual *use* of spoken English by the speaker in social interaction.

When we are able to view speaking/pronunciation skills from these multiple perspectives, we feel that the student can focus on one aspect of his speaking without neglecting or denying the other facets of his speech. In this way, he can concentrate his efforts, for example, on the particle perspective and work on self-monitoring skills that involve the establishment of an auditory/kinesthetic basis for speech production and perception. This goal of self-monitoring is made *explicit* for the ESL student in material recently developed by Joan Morley at the English Language Institute, The University of Michigan, *Improving Spoken English—An Intensive Personalized Program in Perception/Pronunciation/Practice in Context*[2] (to appear 1978):

> To improve your pronunciation of English, you need to learn to monitor your own pronunciation by consciously listening to yourself, and feeling the movements of your lips, tongue, and jaw. Gradually you will recognize the sounds which are not quite correct. Self-monitoring is the foundation of self-correction.
>
> You will be *watching* the movements of the lips, of the lower jaw, and of the tongue, and *feeling* movements of the lips, of the lower jaw, and of the tongue. In addition you will be *listening* to the changes in vowel sounds which you can produce as you move the lips, the lower jaw, and the tongue. These, then, are

---

[2] The passage cited here is a transcription from the videotape that accompanies Joan Morley's *Improving Spoken English* rather than from the text itself.

the three most important self-monitoring skills: *WATCHING, FEELING,* and *LISTENING.*

Tagmemics would support Morley's position. Pike claims that one way of developing these self-monitoring skills is by establishing an articulatory *flexibility* that will allow the student to make a single modification that will have an effect on his *entire* sound system. This, of course, is no small task as it involves Guiora's notion of "ego boundaries". Guiora's research at the School of Medicine, The University of Michigan, has led him to postulate that the lowering of ego boundaries has direct significance for the student's ability to attain native-like pronunciation in a foreign language. Here, however, we are specifically interested in the implications of Pike's beliefs concerning flexibility for the student of English as a Second Language. Pike (1977)[3] claims that the first type of flexibility involves the division of the phonological space in the mouth. Through the use of drills that enable the ESL student to divide and sub-divide *ad infinitum* the phonological space from front to back [i——→ɯ] or high to low [i——→a], the student will have available to him all those sound possibilities that occur in between the two articulatory extremes. This type of drill is particularly useful for students whose native language has only a limited vowel inventory.

The second type of flexibility that Pike stresses involves the modification of the phonological grid *as a whole*. By making a *single* change—such as shifting the entire grid either forward or back in the mouth—*all* sounds are affected. Drills that teach modification of the phonological grid as a whole—larynx high or low, tongue root forward or back, lips spread or rounded, velum open or closed, etc.—provide the ESL student with an inventory of infinite phonological possibilities.

Pike concludes (1977) that, "all of these kinds of things put together give you the flexibility which you need. Now this flexibility is above all what you need if you're going to start to learn a language."

As we have already observed, one of the basic components of tagmemics is the notion of *variation*. Tagmemic theory isolates at least three sources of phonological variation that have direct significance for the teacher of English as a Second Language. They are: (1) speed of utterance, (2) placement in a stress group (or other high-level phonological unit), and (3) spread of phonetic features.

Sound changes that occur due to the speed of the utterance have been referred to in the literature as "fast speech phenomena." Illustrations include Pike's '*Did you enjoy it?*'——→[ǰɪnǰɔ́iɪt], '*Let us go eat*'——→[skuit], and '*Ed had edited it*'——→[ɛdədɛ́dətədɪt]. Allophonic variations of this sort are also found in John Morley's new material (to appear 1978). Specific examples include:

---

[3] This discussion of flexibility is paraphrased from the first film "*Voices at Work*" in Pike's television series ***Pike on Language.***

1. *'Can't you come or don't you want to?'* ——→ [kǽtʃkəm ɚdṍtʃwaɾ̃ə]
2. *'When's her birthday?'* ——→ [wínzəbɝ̀θdei]
3. *'It's time to go.'* ——→ [stáɪmrəgoʷ]

Related examples include forms that vary due to degree of carefulness—i.e. "casual speech." Arnold Zwicky (1972:608) separates casual speech processes into two categories, "serving either ease—assimilation, neutralization, insertion of transitional sounds—or brevity—simplification of geminates, vowel contraction, deletion of weakly articulated segments, monophthongization."

The teacher of pronunciation must be aware of the fact that these forms are not limited in occurrence to so called fast speech situations. Rather, these phonological processes are often the norm even in formal situations. Failure to comprehend them puts the foreign learner at a considerable disadvantage.

Phonological variation can be due to the placement of a sound in a stress group (or other phonological unit). Thus, the phoneme /æ/ in the isolated form *'an'* becomes /ə/ in a larger phonological unit, as in *'He put an apple on the plate'* (Pike and Pike, 1977:9.2). Or, again /æz/goes to /æʒ/ in *'as you'*. This relation of sounds to their larger embedding unit illustrates the central role of *hierarchy* in tagmemic theory.

Spread of phonetic features may also be the source of phonological variation in morpheme variants. Progressive or regressive *assimilation*—the process of one sound becoming more like its neighboring sound—is one of the most common processes of phonological variation throughout all known languages. An obvious example in English is in the phrase *"won't you"* where the voiceless stop consonant [t] becomes the affricate [tʃ] through the process of assimilation.

Another aspect of tagmemics that has direct significance for language instruction is in the *presentation* of classroom material. Tagmemics makes use of the notion of *paradigmatic drills*. This notion and its impact for language teaching is explored in detail by Ruth Brend (1967, p. 33-35) in her article "The Return of the Paradigm". The following is a *citation paradigm* of plural formation in English:

	[s]		[z]		[iz]
cat	-s	cab	-s	dish	-es
cup	-s	dog	-s	glass	-es
cook	-s	tub	-s	watch	-es
rope	-s	rag	-s	kiss	-es
hut	-s	hood	-s	match	-es

There are times when it is more useful to present information in a *matrix display*. A matrix representation is a multi-dimensional presentation of the significant *relationships* within the data. Regular past tense formation in English is presented in Figure 1.

Interpreting the matrix display from left to right, the top horizontal plane represents the first step in the analysis of the pronunciation of the regular past

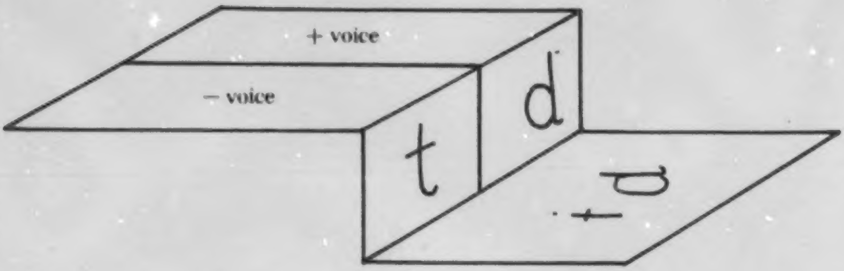


Figure 1. N-dimensional matrix representation.

tense verb ending in English. If the verb stem ends in a voiceless sound, then the regular past ending is pronounced as the voiceless segment [t]. If the last sound of the verb stem is voiced, then the regular past tense ending is pronounced as the voiced segment [d]. These two pronunciations are represented by the vertical plane in the center of the figure. However, if the verb stem ends in a [t] or [d] sound, then the regular past ending is pronounced as an extra syllable [id]. This pattern is reflected in the bottom horizontal plane.

Tagmemic Principles Applied to Error Analysis

Instruction that isolates pronunciation skills from the larger communicative act would follow a different approach in the analysis of a student's speaking errors than we are proposing here. *Rather* than just listing or pointing out a student's errors, *rather* than simply explaining that these mistakes aren't acceptable for various reasons, and *rather* than merely suggesting improvement solely along articulatory lines, we propose to point out the relevance of the student's speaking errors in relation to the larger communicative *event*.

Correction of this type involves more than just one-way communication from teacher to student. If the student is to see the consequences of his mistakes in terms of their *impact* in social interaction, then he must have feedback, not only from the teacher but also from his audience (or peer group). Error analysis that is based upon tagmemic principles, then, must involve interaction between teacher, student and classmates.

Tagmemics suggests the use of various questions for error analysis. They would be used not only by the teacher in helping the students analyze their mistakes, but by the student himself as he evaluates his own mistakes and as he prepares his material. Even during preparation, the student must consider the impact he will have on his audience. Furthermore, as a student's classmates listen to his oral presentation they too would refer to questions like the ones we suggest here, constantly judging the performance in terms of social interaction.

We illustrate this type of error analysis by drawing examples from a student speech debating the topic "The Role of the United States in World Affairs". A transcription of the last section of the text (recorded on videotape) can be found in Appendix I.[4]

1. Were there any mistakes that could have led to "misunderstandings" or breakdowns in social interaction? If there were any difficulties, where did these difficulties occur? Three general tendencies in pronunciation surface in this example that *could* lead to possible misunderstandings. They are:

(1) the simplification of consonant clusters
(2) the substitution of English sounds with Spanish sounds
(3) misplaced stress

Mistakes of this sort are responsible for the student having a "heavy accent" and, at times, precipitating a total breakdown in communication, even though the content of his speech was well organized. Some mistakes that could lead to a misunderstanding in other situations are:

(1) substitutions of the American English voiceless interdental fricative [θ] with the voiceless sibilant fricative [s]. [sɪŋ] or [sɪŋk] for [θɪŋk] is common throughout.
(2) substitution of the American English voiced apicoalveolar stop [d] (or the Spanish equivalent which is dental rather than alveolar) for the voiced interdental fricative [ð]. A few examples include: [dei] for [ðei]; [də] for [ðə]; [dæts] for [ðæts]
(3) substitution of the American English voiced labiodental fricative [v] with the Spanish voiced bilabial approximant [β]. [nέrβus] for [nɚvəs]; and [βɛr] for [vɛri].
(4) misplaced stress: [indóstri] for [índəstri]; [áidiə] for [aidíə]; and [kánal] for [kənǽl].
(5) errors in vowel quality: [ólso] for [álso]; [kóntri] for [kəntrɪ]; [næšən] for [néšən]; [bikóz] for [bikɔ́z]; [stúdiɪŋ] for [stə́diɪŋ]; [kom] for [kəm]; and [obtáin] for [əbtéin].

2. Were there any mistakes in ORAL GRAMMAR? That is: (1) Did any of the pronunciation errors have a grammatical consequence which could have lead to misunderstanding? Or, (2) Were there mistakes that can be attributed mainly to grammatical rather than phonological sources?

Mistakes of both types are everywhere in the text. They are not only "bothersome" for the listener, but can be the source of serious misunderstanding. Just a few examples include:

(1) '*brainstorming*' for '*brainwashing*'
(2) the constant use of '*underdevelopMENT countries*' for '*underdevelopED countries*'

---

[4] Due to limitations of space, only a portion of the text is given here. Most examples cited can be found in this transcription. However, for purposes of illustration, examples from other portions of the speech—not found in this transcription—have also been included.

(3) *'peoples'* which was understood as a possessive, rather than the ungrammatical "double plural"

3. *What ATTITUDES or FEELINGS were conveyed?* This is an especially important question. The student is able to discover what his audience felt he was trying to say, as opposed to what the student himself *thought* he was saying. The disparity is often tremendous. Here, the ESL instructor should especially consider prosodic cues (intonation and voice quality, etc.) as well as non-verbal behavior.

4. *Was this speech APPROPRIATE given the specific social setting?* This leads to a discussion of register and style. While it is not possible to illustrate this point here, it might benoted that the audience felt that this particular presentation, although quite convincing for its content, was entirely too informal for a debate. This was evident in the overall "tone" of the address and also in certain lexical choices. For example, his frequent use of the phrase *'my friends'* seemed a little out of place.

As must be evident from the above suggested procedure for error analysis based on tagmemic principles, a speaking activity is *much* more than an opportunity for a student to "practice" his spoken English. It is a chance for him to *USE* spoken language in *purposive social interaction.*

Tagmemic Principles Applied to Materials Development— The Use of Games via Impact

We have adapted the television games PASSWORD and $20,000 PYRAMID for use in our speaking classes and have also made our games relevant for students entering scientific and technical fields. We also used simulation games and developed one game specifically for engineering students, based on an actual case involving a local contracting company and a utility company.

Games provide a learning environment which encourages social interaction. In tagmemics, meaning is defined as the impact of a language unit upon the listener (together with the form). The gaming sessions demonstrate for the students the consequences of their interactive behavior.

Students also have a *purpose* for correct pronunciation—winning the game; and they are constantly receiving feedback from the other players, students and teacher, who are monitoring the pronunciation. Negative feedback from his gaming partner, for example, would indicate that the student's clue was disfunctional. The student might repeat his clue, modifying it, in order to hopefully achieve the desired consequences—a correct guess from his partner.

Beck and Monroe (1969), in an article called "Some Dimensions of Stimulation", point out that the consequences of the players' course of action during a game are observable. In order to emphasize this aspect of gaming, we videotaped many of the television and simulation games. These video tapes used along with Morley's self-monitoring and self-correcting procedures (see pp. 84-85) and our proposed error analysis (see pp. 87-89), were extremely useful for student and teacher evaluation of speaking activities.

Following are three examples of PASSWORD (see Figure 2). The object of the game is to get your partner to guess the correct word using only one-word clues.[8]

In the first game, the simplification of the final consonant cluster in *'lamp'* to [læm] causes misunderstanding, and leads speaker D to ask "How do you spell it?" In the second game, faulty perception of the final voiced sound in *'eyes'* (heard as *'ice'*) leads to an incorrect guess—*'iceberg'*. In the third game, student C was not sure of the pronunciation of the final consonant in *'cloth'*. This problem, deciding on voiced vs. voiceless sound, led to an additional problem of vowel quality and length, which led to a misunderstanding (and the other team eventually won the game).

"LIGHT"

A: 'Bulb'
B: What?
A: 'Bulb' b-u-l-b.
B: I don't know.
C: 'Lam(b)'
D: 'Lamp'?
C: 'Lam(b)'
D: How do you spell it?
C: L-a-m-p. 'Lam(b)'
D: (Pointing to the light) I don't know the name.
A: 'Sun'
B: 'Light'.
A: Right!

"GLASSES"

A: 'Eyes'
B: 'Iceberg'
C: 'Short-sighted' 'Short-sighted'
D: 'Short-sighted'? 'Winter' I don't understand the word.
C: It's too difficult for you.
A: 'Eyes'
B: 'Ice' . . . 'Glasses'. What a brain! Oh my goodness!

"CURTAINS"

A: 'Window'
B: 'Mirror'?
C: 'Cloth' (pronounced as [klɔ:ð]) [klɔð]. [klaθ]? [klɔð]. [klaθ]?
Teacher: Yes.
C: 'Cloth'
D: 'Open' (In response to 'close') No, pass.
A: 'Window'
Teacher: What did you say? 'Open'?
B: Yeah.
C: 'Blind-uh' (for 'blinder')
D: 'Blind-uh'? Pass.
A: 'Cover' 'Cover'
B: Yeah . . . .

. . .

B: I think I know the word. . . .
Is the word . . . 'curtain'? curtain'?
Teacher: Yes.

Figure 2. Password

Another game we use that allows the students to observe the consequences of their verbal behavior is adapted from the television game $20,000 PYRAMID. The object of this game is to get your partner to guess seven items (words or phrases) which belong to the same category. A sample category might be "Things that are white", and the seven items to be guessed could include: snow, milk, chalk, a nurse's uniform, a wedding gown, and paper. In contrast to the one-word clues in PASSWORD, the students can use gestures, short phrases or sentences, etc. as clues. In LANGUAGE PYRAMID, we are no longer interested in viewing language from the static perspective, as in PASSWORD but rather we are interested in the flow of speech, a dynamic moving entity.

---

[8] Both PASSWORD and $20,000 PYRAMID are subject to time limitations that the teacher can vary in accord with the students' ability.

The only example of a static particle that blocks the flow of the game describing "Things that are frightening" is *'imaginic'* (for *'imaginary'*), which is understood to mean *'mountains'*. The following are two examples of our LANGUAGE PYRAMID (see Figure 3).

"THINGS THAT ARE TYPICALLY AMERICAN"

Teacher: Ready? Begin!
A: An American food. Very long American food.
B: A "hot dog".
A: A game. You take the ball with you and you go (run over/to) the goal.
B: Ah! Basketball.
A: No. Ah . . . the ball over here.
B: Football.
A: Right! A city in the United States where there are many actors and go to the movies the actors. In California . . . I think.
B: Hollywood.
A: Right. The President of the United States.
B: Carter.
A: A person who ride a horse and are very typical in the U.S. and they with the horse they . . . they push the cows.
B: Oh. Cowboy.
A: An animal who is going to be extinct in the . . .
B: Bison.
A: No.
B: Buffalo.
A: The other one. Something that there is in the sky and is very bright. Is very bright, very bright and [OVER] there are many bright points and. . . .

Answers: *1. 'hot dog' 2. 'football' 3. 'Hollywood' 4. 'Jimmy Carter' 5. 'cowboy' 6. 'buffalo'* 7. *'Stars and Stripes'*

"THINGS THAT ARE FRIGHTENING"

A: Exam that you make at the end of semester, 150 or 160.
B: TOFEL. Oh! Michigan Test!
A: Animal. Very big.
B: Bear.
A: No. It's imaginic.
B: Oh, in the mountains?
A: No. It's imaginic, the name of the animal, the big animal . . .
B: Monster.
A: Yes! One animal. Poison. Without hands.
B: Snake!
A: Yes. American film. A big monster.
B: "King Kong"
A: Fighting between two governments.
B: War.
A: Uh . . . it's a storm. In the United States. It's very hard storm.
B: Hurricane.
A: No another one.
B: Tornado.
A: When you buy one book, you have to pay . . .
B: TAX! !

Answers: *1. 'The Michigan Test'* 2. *'monster'* 3. *'snake'* 4. *"King Kong"* 5. *'war'* 6. *'tornado'* 7. *'taxes'*

Figure 3. Language Pyramid

Conclusion

It has been our assumption that instruction which: (1) enters into a discussion of language on the level of social interaction, and (2) treats meaning as impact comes closer to fulfilling the needs of our students than do more traditional approaches. We have attempted, therefore, to outline some basic concepts of tagmemics, such as impact, perspective, hierarchy, and variation, and to examine their applications to the teaching of ESL speaking skills. Specific applications of tagmemics discussed include: presentation of material, error analysis, use of games, and development of self-monitoring skills.

REFERENCES

Beck, Isabel, and Bruce Monroe. 1969. Some dimensions of simulation. *Educational Technology*, October, 45-49.

Brend, Ruth. 1967. The return of the paradigm. *Language Learning 17*, 33–35.

———. 1976. *Applied tagmemics*. Unpublished manuscript.

Catford, J. C. 1967. Intelligibility. In W. R. Lee (ed.), *English Language Teaching Selections 2*, London, Oxford University Press, 142–150.

Duke, Richard D. 1974. *Gaming: the future's language*. New York, Sage Publications, distributed by Halsted Press, a division of John Wiley and Sons, Inc.

Guiora, Alexander Z. *et al.* 1970. The effects of experimentally induced changes in ego states . . . on pronunciation ability in a second language: an exploratory study. The University of Michigan, School of Medicine. Bibliographical references incomplete.

Morley, Joan. 1978. (in press) *Improving spoken English—an intensive personalized program in perception/pronunciation/practice in context*. Ann Arbor, The University of Michigan Press.

Pike, Kenneth L. 1977. *Pike on language*. Parts 1 (*Voices at work*) and 3 (*Waves of change*). Ann Arbor, University of Michigan Television Center. ¾ inch videocassettes.

Pike, Kenneth L. and Evelyn G. 1977. (in press) *Grammatical analysis*. Huntington Beach, California, Summer Institute of Linguistics.

Troyka, Lynn Quitman. 1973. A study of the effect of simulation-gaming on the expository prose competence of college remedial English composition students. New York University, Ph.D. dissertation.

——— and Jerrold Nudleman. 1975. *Taking action: writing, reading, speaking, and listening through SIMULATION GAMES*. Englewood Cliffs, New Jersey, Prentice-Hall, Inc.

University of Michigan Extension Gaming Service. 412 Maynard, Ann Arbor, Michigan 48109.

Zwicky, Arnold M. 1972. On casual speech. In *Proceedings from the Eighth Annual Regional Meeting of the Chicago Linguistic Society*. 607-615.

APPENDIX I

THE ROLE OF THE UNITED STATES IN WORLD AFFAIRS (SELECTIONS)

Maybe you can ask me another question. If it is your mean, why do Latin Americans
[kwɛsion]
come to United States, why do, why do, why you stay in United States? Well, I guess
[kom]
we like North Americans people. I guess we like United States but we don't like the
[do:n]
politicians who intend to produce a brainstorming in Latin, North Americans people. This
[politísiɔns] [brenstórmɪŋ] [nort] [dis]
is the reason because United States, North Americans people have this rebellions every day.
[dɔ] [bikóz] [nort] [hæf]
They are not agreed with Viet Nam, or with Panama Canal Zone and they think, North
[wits] [raŋ] [kánal] [siŋk]
Americans people, they think that this situation have to change but the govern(ment)
[θiŋk] [hæf]
don't like and don't intend to change the situation, ¿no? Well, I guess the United States
[do:n]
can help and we also need a help but the help not for obtain all countries, all industries or
[ólso] [obtáɪn] [indóstris]
all people. Its only help technology and this is a reason because we are here in the United
[bikóz] [heɚ]
States, Latin Americans people, studying, learning English and improving our technology.
[stúdiɪŋ]

And I guess if when we return to our countries we can have a good, a good idea about
[ritúrn] [æ] [áidiə]
United States and I guess and also good studies and I guess no Latin Americans people
[god] [stódis]
can speak with good words about the United States, about they students, they people, but
[wits] [dei] [dei]
no about they politicians. Okay, that's all.
[dei] [politísiɔns] [dæts]

A New Direction for Language Lab Programs[1]

Wu Yi So and Doris L. Scarlett
Michigan State University

This paper focuses on the potential of the language laboratory for ESL instruction. In Part I the authors describe a language lab program for advanced level students. These newly developed materials represent a synthesized approach to second language learning and reflect an earnest attempt to bridge the gap between English as it is presented in the ESL classroom and as it is used in American university lecture halls and discussion groups. In Part II the authors suggest possibilities for adapting the same rationale and format for other purposes and particular needs. They stress the importance of the ESL teacher's involvement in materials development and offer guidelines for those who might undertake such a project.

Many ESL teachers and administrators can probably recall the heyday of the language lab boom in the early and mid-sixties. For some, it was soon followed by a period of disappointment and discouragement. People began to doubt that machines could really facilitate language learning; in many cases, the expensive electronic gadgetry brought more frustration than help.

The English Language Center at Michigan State faced the same problems and set about to re-assess its language lab program. Careful study and analysis convinced us it was the software, not the hardware, that needed attention.

After deciding to explore ways and means of putting our lab facilities to better use, we began to revise our old tapes and make new ones. An earlier paper (So, 1974) related the rationale and general direction of a new language lab program for advanced students. In the last three years we have concentrated our time and effort on refining and further expanding this series of tapes, which reflects an earnest attempt to bridge the gap between the English our students learn in the language classroom and the English they need to know and use in lecture halls and discussion groups when they take academic courses. The approach we have used is what H. Douglas Brown describes as "cautious, enlighted eclecticism" (1975, p. 82) or what James Ney has called a "synthesized" approach (1973). We have attempted to combine the best of the audio-lingual habit theory with the cognitive-code learning theory.

In Part I of this paper, we describe in some detail the lab materials we have developed; in Part II we suggest the possibility of adapting the same rationale and format for other purposes and particular needs, and we offer guidelines for those who decide to undertake such a project.

---

[1] This paper, together with sample tapes, was presented as a demonstration at the 1977 TESOL Convention in Miami. The original paper has been revised and expanded for publication.

PART I: A NEW LAB PROGRAM FOR ADVANCED ESL STUDENTS

Description of the Tapes

Our advanced level language lab program consists of over 50 lessons arranged in 26 units which are sequenced according to difficulty but also with an eye and an ear to topic, format, and style. A few of the units consist of only one lesson, designed to be completed in one 50-minute class period; others involve three lessons; the majority are two-lesson units. Regardless of the number of lessons, each unit consists of a tape-recorded dialogue, speech, or lecture, followed by a series of language-oriented exercises based specifically on that particular text.

For each unit, the student is provided with worksheets, a printed text of the speech/dialogue/lecture, and sometimes a glossary or other hand-outs. The teacher also has these, plus a tape-recording of the text and exercises and a teacher's script which is an exact transcription of the tape. The script for each lesson also includes instructions for the teacher, and it is accompanied by a cover sheet which gives an approximate running time for each exercise on the tape.

A dialogue, speech, or lecture is the focus of each unit. This "text" serves a three-fold purpose. It is, first of all, an example of *real communication* in English—a speech by a U.S. President, a dialogue between two students, a lecture by a professor. Second, each text provides *information* about any one of a number of topics.[2] A third and very important purpose of each text is to provide a *context* for developing various language skills.

In addition to covering a wide range of topics, the texts serve to familiarize students with formal as well as informal speech. For example, there are speeches by Gerald Ford, Richard Nixon, John F. Kennedy, and Martin Luther King. There are dialogues between native Americans and between a native American and a fluent foreign student. And there are lectures by MSU professors and English Language Center teachers. The lectures are recorded on tape by the contributors who wrote them, thus exposing the foreign students to a variety of American voices, dialects, and—just as important—the various speaking styles found in American university classrooms.

A uniform format is followed throughout the entire 26 taped units. Each begins with a recorded text, followed by a vocabulary exercise, and then other exercises focusing on various receptive and productive skills. Although the format is uniform throughout the series, the same kinds of exercises are not given in each and every lesson, because some texts do not lend themselves to certain exercises. Within each skill area, the exercises are cumulative and follow a standard format. Also, the activities are sequenced from the relatively easy to the more difficult. All exercises are designed to encourage *active* student participation at all times.

---

[2] A sampling of unit titles indicates the variety of topics covered: Rodgers and Hammerstein, Advertising in the U.S., American Dialects, American Country Music, Note-Taking, Outlining, American College Dictionaries, Black Leadership Goals, Impact of Technology on the Popular Arts, Chinese Acupuncture, Oil, The American Civil War, Soybeans, Flight of Apollo 11, Improvement of Nutrition in Food Crops, Martin Luther King's "I Have a Dream" Speech.

A typical unit does not actually begin in the language lab. Most units have handouts—glossaries, charts, maps, outlines, etc. Students are encouraged to read the handout before coming to class because it provides background information. For example, a glossary, with items that are generally technical, cultural, or historical in nature, helps the student to better understand the recorded text. Or the handout may be a map or outline the lecturer refers to; the student should have it in front of him as he listens to the tape in lab class.

Once students are in the language lab class, they first listen to the recorded text—*without* having read it before and *without* reading it as they listen. Students are encouraged to get as much of the meaning as possible during this first hearing of the text. The first exercise following the text is almost always a vocabulary exercise, which focuses on words in the text with which the student is likely to be unfamiliar or words which, while perhaps familiar, may not be part of his active vocabulary. Various other exercises follow the initial vocabulary exercise.

Certain guidelines have been followed in the sequencing of exercises within a particular lesson. Aural comprehension exercises always come *after* vocabulary exercises, so students will have the benefit of the vocabulary exercises. Exercises which deal primarily with listening skills are alternated with those which require more oral production—to give variety in student activity during the class period (Rivers, 1968, p. 338). Exercises which require a significant amount of writing and are to be turned in to the teacher usually come toward the end of the lesson, in order to accommodate individual learning pace; dictation, notetaking, and outlining exercises are examples. The total time for the lesson is usually kept between 35 and 40 minutes, thus allowing sufficient time for the teacher to check performance and answer questions after each exercise and still making it feasible to fit the entire lesson into a 50-minute class period.

Content of the Exercises

Vocabulary. Since our main objective is to better prepare our students to do academic work, we have incorporated in our material the development of a variety of language skills and study skills. One of these is the skill of guessing meaning from context. In every lesson there is a vocabulary exercise focusing on words we expect students to learn.[3] In general, we select a word on the basis of how important the word is to the foreign student's understanding of the text. Sometimes it is impossible to introduce all the possible new words in a text, so we try to select those the student is likely to come across again in his speaking and listening.

---

[3] Words listed in the *glossary* are words students need to look over and understand, but not necessarily learn. We ask students to look over a glossary *before* coming to lab class, but we tell them *not* to study a vocabulary exercise ahead of time; we want them to actually go through the experience of figuring out the meaning of a word (or phrase) without previous preparation.

The ultimate goal for all our vocabulary exercises is not teaching vocabulary *per se*, important though this may be; rather, it is to help the student develop the habit of guessing meaning from context and to help him build up his confidence in making intelligent guesses. In each vocabulary exercise, we first ask the student to listen to the word and the sentence in which the word is used. (Both the word and the sentence appear on the student worksheet.) Then there is a pause on the tape, allowing time for the student to read the three choices given on the worksheet and circle the correct meaning. We encourage the student to guess even though he may not be sure. After a group of six to ten words has been presented in this way, the correct answers are given on tape. The student gets immediate reinforcement by repeating the word and the sentence after checking each answer.

Throughout the lessons we try to impress upon the student that a word can sometimes have more than one meaning and that he has to choose the meaning that best fits the sentence. Sometimes we purposely use other possible meanings as choices to force the student to select the meaning that is most appropriate. At other times, we use as choices words that have already been introduced in previous lessons, to provide reinforcement.

Aural Comprehension. Every unit in our series of advanced level tapes includes at least one, and often two or three, aural comprehension exercises. These involve the extraction of specific information from the text. Actually, we work on aural comprehension skills at two levels.

The first time our students hear a recorded text—at the beginning of a unit—they are in a situation similar to that of students in university lecture classes. For this reason, we ask them *not* to read the text ahead of time. (We point out that when they begin taking academic courses their professors are not likely to hand out texts and read the lectures word-for-word!) We encourage students to listen, this first time, for general understanding, for the main points—even though they may not understand every word the speaker says.

After the students have heard the text and done the vocabulary exercises and perhaps some other exercises, we present an aural comprehension exercise. This may involve true/false, multiple-choice, completion, or matching items, or perhaps short-answer or reflective questions. Regardless of the type of items, the student is always given a minute or two to look over the exercise before listening to the text again. As he listens, he marks his worksheet. For most exercises, the answers are given immediately after the replay of the text. However, if the exercise involves matching items or some type of questions, the teacher may stop the tape after the replay of the text and give students a little extra time to complete the exercise. If reflective questions are involved, the instructor collects the worksheets, checks the exercise, and returns the papers at the next class meeting.

Pronunciation and Intonation. Many units include a pronunciation exercise focusing on the primary stress in words. Both listening and speaking activities are involved in this type of exercise. The words included in such a drill always come from the text and are selected on the basis of our observation of the actual needs of our students.

First, the student listens as a voice on tape gives the correct and incorrect pronunciations for each word in the exercise. In the beginning lessons of the series, exercises are restricted to 8-10 words; later on, up to 15 words per exercise are included. A few sample items are listed below:

1.	A	ássignment	B	assígnment
2.	A	cónsultant	B	consúltant
3.	A	supérior	B	súperior
4.	A	compónents	B	cómponents
5.	A	índustrialized	B	indústrialized

Following along on his worksheet (where each word is typed twice, once with the stress correctly marked and once with it incorrectly marked), the student listens and circles the pronunciation which he thinks is the correct one. Next, the answers are given on tape; after each answer, the student is asked to repeat the correct pronunciation. As the last step in the process, each word is correctly pronounced one more time; while the student listens, he inserts a stress mark in the appropriate place (on a new, unmarked list of the words) and repeats the correct pronunciation again. While some may feel it is wrong to present incorrect pronunciations, we have found it helps the student learn to recognize his errors and to be discriminating.

Another type of exercise involves intonation patterns, specifically the intonation patterns of tag questions. Here, too, the focus is first on receptive and then on productive skills. First, the student listens to brief explanations of falling and rising intonation, including the purpose and meaning of each and several examples of each from the text. Second, the student listens to a series of statements with tag questions; these are printed on the student worksheet. As each statement is read twice on tape, the student must decide if he hears rising or falling intonation and mark the statement on his worksheet with an appropriate arrow—up or down. Third, the answers are given, and each statement is repeated by the student. Teacher monitoring is extremely important at this point in the exercise. If only a few students are experiencing difficulties (in their identifying or producing the two intonation patterns), the instructor may suggest they repeat the exercise during "open hours" or meet with the instructor after class. If, however, a majority of the students express difficulty, the instructor may choose to rewind the master tape and repeat the entire exercise.

Oral Production. The student is led to gradually become familiar with the rhythm and the intonation patterns of spoken American English through three stages of oral production:

(1) In the first few lessons we ask the student to do repetition drills. A

native speaker reads a selected paragraph or two of the text, breaking up the sentences into natural groupings and leaving a long enough pause between word groups so the student has time to repeat. After the student has finished repeating the entire paragraph, he is asked to rewind his tape and listen to his voice, comparing his pronunciation and intonation to that of the native speaker. Playback time is a good time for the instructor to help individual students and perhaps discuss specific pronunciation or intonation problems on an individual basis. In order to take advantage of such opportunities, we encourage the lab teachers to keep a monitoring chart for each student.

(2) The second stage is the reading fluency drill. The student reads along, pacing the speaker rather than repeating after him. The student is led to approach this task step by step. First, we ask the student to consciously recognize how a native speaker groups his words together. As the student listens to the reading of a passage, he is asked to put down slashes between words whenever he hears a slight pause. In later lessons a reading fluency drill also includes listening to the passage a second time, concentrating on important words—words which are spoken louder and with greater emphasis. The student is asked to underline the important words as he listens. After he has put down the slashes to indicate the grouping of words and underlined the important words which require emphasis, then he is ready to read along, pacing the speaker and imitating the speaker's pronunciation and intonation.

(3) The third stage is an oral composition which the student records on tape. There are three such exercises spaced throughout our 26 units. In each case the student is asked to summarize a discussion or comment on a topic he has just studied in lab class. He is encouraged to express his ideas freely, using the vocabulary he has just learned. The oral composition does not have to be very long. The first time, we ask the student to make a one or two-minute recording; each succeeding oral composition is longer. Often, a student will record two or three versions and then select the best one for the teacher to listen to. After collecting and listening to all the taped oral compositions, the instructor schedules a conference with each student so that they can listen to the recording together, talk about error types (both in phonology and syntax), and discuss ways for improvement. We realize that these student-teacher conferences are rather time-consuming, but we are convinced that they are beneficial to both student and teacher. The teacher gets to know each student better and becomes more aware of individual learning styles, strategies, and problems. The student benefits from the individualized instruction and attention.

Dictation. Another skill we can help the student develop in the lab is that of encoding utterances into written symbols—what is traditionally known as "taking dictation." Cumulative practice of this kind, from short sentences to longer ones, from simple structures to complex ones, helps the student improve in the mechanics of writing (spelling, punctuation, capitalization, contraction, etc.). And student performance in a dictation exercise may be an indicator of overall proficiency or a clue to certain language problems.

We want the dictation exercise to be a learning experience for every student. Slow students sometimes get frustrated and discouraged if they don't have enough time to write; therefore, we try to pause long enough between sentences so that the slow students will not be discouraged and the fast ones can still feel some challenge. We generally read each sentence two times; if a sentence is significantly longer than what the students are used to, it is read three times. The length of the pause between the first and second readings is determined by asking a native speaker to write the sentence twice. We have found that if a native speaker has enough time to write the sentence twice, then the pause should be long enough for the foreign student to write the sentence at least once. During and after the second reading, the student can fill in what he missed the first time.

Note-Taking and Outlining. Because the majority of our students at the English Language Center are learning English in order to pursue a course of study at some American university, we feel that helping them to develop their note-taking and outlining skills is a very important goal. We build the two skills on a base of information provided by two lectures specifically designed for us by Dr. Robert Geist, an MSU faculty member. In the first lecture, Dr. Geist discusses note-taking, makes specific suggestions for taking notes in English, and introduces the student to some conventional abbreviations. In the second lecture, he discusses outlining—illustrating all of his remarks with examples from his previous lecture on note-taking. The Outlining unit includes a handout which gives both a topic outline and a sentence outline for the lecture on note-taking.

Actual skills development begins on a very small scale: students are asked to listen to about two minutes of dialogue from the American Country Music tape and fill in a single detail after each of the six key words on the worksheet. This is obviously very easy, and the students are usually quite successful. Each subsequent note-taking exercise is in some way more difficult. Also, as the note-taking exercise itself becomes harder, it is also integrated with outlining. For example, in one exercise the student take notes by filling in a skeleton outline for one major segment of a lecture. As this type of note-taking/outlining exercise progresses through our series of units, the outline covers more of the lecture or becomes more complex, or the student is provided with less information on the worksheet. The final exercise is a combination of note-taking and outlining: the student listens to an entire lecture and takes notes in outline form.

Useful Expressions. Included throughout the entire series of tapes are exercises which introduce the student to idiomatic expressions, multiple-word verbs, and preposition combinations.

We have developed a special format for the treatment of idioms and multiple-word verbs. First, the student listens to a definition of the expression; second, two or three examples of usage are cited and then repeated by the student—the first example always coming from the text; third, the teacher stops the tape while students fill in the blanks in a set of sentences, using the expressions just studied; then, the teacher begins the tape again, and the students hear the correct answers and repeat both the expression and the sentence in which it is used; fourth, for homework, the student is asked to write sentences of his own using the expressions.

In exercises which focus on preposition combinations, the student first hears a series of sentences from the text which contain the preposition combinations to be studied. These sentences are printed on the student worksheet, but the preposition is omitted in each case. The student must listen to the sentence as it is given on tape and write down the preposition he hears. After all the sentences have been read, the correct answers are given and the student is asked to repeat each preposition combination.[4] After the students have finished this first part of the preposition exercise, and heard and repeated all the correct answers, the instructor stops the tape while students fill in the blanks in another set of sentences with appropriate prepositions. As the answers are then given on tape, the student is asked to repeat the preposition combination and the completed sentence.

Conclusions

(1) Responses from both our students and teachers have been very favorable and indicate that the format goes a long way toward better meeting the needs of our university-bound students. We hope to be able to objectively measure the effectiveness of this material and have some empirical research data to report in the near future.

(2) The program is most successful when both students and teachers receive adequate orientation. Students must know how to use the machines, what is meant by various instructions on the tapes, and how and when to use the printed materials which accompany the tapes. Teachers must have a working knowledge of both their master console equipment and that used by students; they must know what the machines can and cannot do. They must also understand fully the rationale behind the program and the main objectives not only of the tapes as a whole but also of the cumulative exercises in each skill area.

(3) The teacher-student conferences are extremely vital. They are an important aspect of our approach to the development and evaluation of oral production skills. Just as significant, however, is the opportunity for a one-to-one exchange without the constraints of the language lab setting.

(4) Although these tapes constitute a self-contained program of instruction, it has become increasingly clear that correlation with other parts of our

---

[4] Here is an example of what appears on the *student worksheet:*

1. Before the Civil War, advertising consisted________small paragraphs.
2. Benjamin Franklin emphasized the rewards derived________using a product.

The *teacher's script* for the above example would indicate that the following is recorded on tape:

1. of
   Repeat: consisted of (Pause)
2. from
   Repeat: derived from (Pause)

total intensive language learning program is both possible and advantageous. Specifically, we have found that these materials can be correlated with the advanced level writing classes. Lab units such as Note-Taking and Outlining lay the groundwork for further skill development in the writing class; topics introduced in lab can be the focus of discussion and written assignments in writing class; the style, organization and methods of development used in lab texts can be analyzed—and criticized—in writing class. These are just a few of the possibilities.

PART II: PRODUCING YOUR OWN LAB MATERIALS

Who

Who is best prepared and equipped to write language lab materials? Who should test and evaluate them? Who can most influence their success?

The language lab materials described in Part I represent many, many hours of writing, revising, recording, editing—and much thought and discussion. Professors, instructors, graduate assistants, and lab technicians contributed various kinds of expertise. But the most crucial element has been the contributions of ESL teachers—their combined intuition, observations, and experience. For obvious reasons, those who are not ESL teachers and have not had previous teaching experience in an ESL language lab setting are at a distinct disadvantage when it comes to writing lab materials for ESL students.

The writing of new lab materials is a major undertaking in itself. But it is only the first step in the development of a workable lab program. The second step is the initial use and evaluation of the materials, for the purpose of answering questions such as the following: Does the information provided by the materials meet the needs of the students? Are the exercises appropriate to the students' level of proficiency? Are the instructions adequate? Is the tape recording clear? Is too much—or too little—expected of the student? of the teacher? Are the exercises sequenced appropriately? Do the written materials coincide with the taped materials? Did the typist omit a line—a vocabulary choice—an exercise? Did the technician forget to splice out a cough—or forget to splice in a replay of the text?

The answers to these and other such questions may be found by a simulated "run-through" of the lesson by the writer and/or editor. Or a classroom teacher may actually use the materials in a real class and report back to the script writer. Ideally, however, the ESL teacher who has written the material and an ESL teacher who is in a position to use the material should *together* try out each lesson in an actual lab class.

The language lab teacher must play an active and very significant role if the completed materials are to be effective. Too often, ESL teachers, both the new and the inexperienced, feel that lab class is an "easy" teaching assignment. Such teachers reason that "all you have to do is turn on the tape—just press a few buttons—you don't even have to draw up a lesson plan—everything is done for you." While it may be true that the language lab lesson is carefully planned and written ahead of time, the lab teacher must be as familiar with the teacher's script as if he had written it himself; he must know when and where to stop the tape; he must judge the extent to which he can alter or adapt the lesson—and whether or not, depending on the situation, this will help or confuse his students; he must develop efficient and effective ways of eliciting student feedback during the limited times when the tape is stopped; he must make quick decisions as questions are asked—should he answer individually or tune in the whole class; he must develop a sensitivity which passes beyond the glass wall that separates him from his students and the earphones and switches which connect them; he must develop excellent monitoring skills and keep accurate records of each student's progress. He must fully realize that one of the most important lines in the teacher's script is the one which reads, "Stop tape. Check performance and answer questions." Above all, he must remember that *he*—not the tape or the tape-recorder—is the teacher.

What

Although we have every confidence in the rationale on which our advanced level lab program is premised, and although we strongly feel that the format of our lab lessons is unique, we still realize that our particular materials may not be appropriate for all types of ESL students and programs. However, we do feel that both the rationale and the format may be adapted according to particular needs and at various levels of ESL instruction.

There are numerous possibilities for adaptation. Here are just two suggestions. (1) Those interested in developing ESP (English for Special Purposes) or EAP (English for Academic Purposes) lab materials might select appropriate base information (lectures, readings, etc.) and then use the exercise formats we have described to develop particular skills within the focus of the ESP or EAP area. (2) Those in need of lab materials for other levels of ESL instruction might adapt the format accordingly. For example, the staff at our English Language Center hopes to be able in the near future to begin work on a set of lab tapes for our intermediate level students. Lesson and exercise formats will be similar to those for our advanced level materials, but instead of basing the materials on speeches, dialogues, and lectures, we will base the intermediate materials on a book of readings for intermediate ESL students.

How

Time and space do not allow us to relate in depth the procedural do's and don't's we have followed in developing our own lab program or the changes we would make in these procedures given the luxury of hind-sight. However, we will attempt a few suggestions:

(1) Plan ahead. You should have in mind, from the beginning, an overall plan as to the purposes, limits, scope, length, sequencing, etc. of the materials

you are building. As you plan, keep in mind the facilities and manpower with which you have to work. As soon as possible, establish basic lesson and exercise formats. In fact, it may prove helpful to set up a typing guide—not only for the typist, but for everyone involved in script writing.

(2) Be prepared to revise. Even with expert planning, revision is inevitable; there are bound to be some problems and situations you couldn't have anticipated. But revise with care and caution. Don't make revisions on an isolated basis. It's better to take the time to make one major revision that will hold up than to make several quick and expedient revisions that don't address the basic, overall problem.

(3) Keep up-to-date records of what you have and have not done. For example: (a) Use a color-coded card file system to keep track of all vocabulary items, idioms, multiple-word verbs, preposition combinations, pronunciation items, etc. treated in the exercises. (You might also include glossary items and vocabulary choices or distractors.) (b) Keep a file folder for each unit or lesson. Include copies of all drafts; when revisions are made, note them—and the reasons for making them and the date the changes were actually made. (You might want to draw up and duplicate a routing chart or check-list to accompany each folder.)

(4) Share ideas at every stage. You need feedback from everyone: teachers, students, technicians, recording artists, typists. Here are some specific procedures we found helpful: (a) Consult several references when writing definitions for vocabulary items, multiple-word verbs, and idiomatic expressions. We regularly consulted the dictionaries and handbooks listed among the references at the end of this article. We found that the *Oxford Dictionary* consistently gave simpler, more concrete definitions, but care must be taken to distinguish differences in American and British usage. (b) Consult at least one other teacher and/or script writer before you finalize a rough draft or a revision. (c) When tapes and printed materials for a particular unit are ready for a trial run, actually use them in a real class. Ask the teacher for comments. Better still, sit in on the class yourself and make note of student and teacher reactions, technical problems, mistakes, misunderstanding or misinterpretation of instructions, performance levels, pauses that are too long or too short, etc. (d) Even when the materials are in "final" form, continue to provide channels for feedback. We have found it helpful to attach blank note paper to the inside of the file folders in which our master teacher's scripts are stored in the lab; thus, teachers can easily write notes about problems, mistakes, etc. which they encounter as they use the tapes.

(5) Choose recording artists with care. While it is hoped that the speeches, dialogues, and lectures will provide a sampling of American dialects and speaking styles, the language execises themselves should be recorded by native speakers whose voices are pleasing and clear *on tape*. (Sometimes an audition is needed.) We prefer to record each lesson with two contrasting voices, one male and one female, so that no one voice dominates the entire tape, so that

voices don't "wear out" before the end of a recording session, so that the tape does not become monotonous to the students, and so that we can signal changes in exercises or parts of exercises with a change in voice. We have also found that it is extremely helpful to have a third person present at each recording session to direct the recording—someone to make sure the script is followed, to make on-the-spot decisions regarding changes, to record such changes in writing on a master script, to evaluate volume level and voice quality, to judge the lengths of pauses.

(6) Enlist the help of lab technicians who have the patience and the technological ability to do the recording and editing. Good technicians can save you many hours of extra work.

(7) Be prepared to edit. First of all, technical flaws and human mistakes are inevitable—and glaring on tape; they cannot be overlooked. Second, timing can be crucial. Pauses which are either too long or too short can bore or frustrate students. Also, an exercise which you estimated to be 4'30" but actually took 9'14" to record is probably going to throw off the timing of the entire lesson; you'll have to shorten the exercise or adjust or omit some other exercise.

(8) Be aware that developing lab materials requires creativity, tedious attention to details, and a sizeable investment of time, energy, and patience. But the experience will be rewarding and the team work can be fun. As Wilga Rivers has pointed out, "The uses to which the laboratory can be put at advanced levels are limited only by the imagination, resourcefulness, and enthusiasm of the teacher" (1968, p. 339).

REFERENCES

American Heritage Dictionary of the English Language. 1976. New College Edition.

Brown, H. Douglas. 1975. The next 25 years: shaping the revolution. In M. K. Burt and H. C. Dulay (eds.), *On TESOL 75: new directions in second language learning, teaching, and bilingual education*. Washington, D.C., TESOL.

Crowell, Thomas Lee, Jr. 1960. *A glossary of phrases with prepositions*, 3rd ed. Englewood Cliffs, N.J., Prentice-Hall International, Inc.

English Language Services. 1964. *The key to English two-word verbs*. New York, Macmillan Publishing Company, Inc.

Heaton, J. B. 1965. *Prepositions and adverbial particles*. Hong Kong, Longman Group Limited.

Makkai, Adam, (ed.). 1975. *A dictionary of American idioms*. Woodbury, N.Y., Barron's Educational Series, Inc.

Ney, James W. 1973. Towards a synthetization of teaching methodologies for TESOL. *TESOL Quarterly* 7, 1, 3–11.

Oxford Advanced Learner's Dictionary of Current English. 1974. New Edition.

Random House College Dictionary. 1975. Revised Edition.

Reeves, George. 1975. *Idioms in action*. Rowley, Mass., Newbury House Publishers, Inc.

Rivers, Wilga M. 1968. *Teaching foreign-language skills*. Chicago, The University of Chicago Press.

So, Wu Yi. 1974. A New Language Lab Program for Advanced Students. *TESOL Quarterly* 8, 3, 293–304.

Webster's New Collegiate Dictionary. 1975.

Whitford, Harold C., and Robert J. Dixson. 1973. *Handbook of American idioms and idiomatic usage*. New York, Regents Publishing Co.

5

The Miscue-ESL Project

Pat Rigg

State University of New York at Albany

This is a preliminary report of some of the results and implications of a study which analyzed the oral reading miscues of four groups of ESL speakers: Arabic, Navajo, Samoan, and Spanish. Results to date suggest 1) there may be universals in the reading process itself; and 2) one's first language does not determine one's reading proficiency in ESL. The 2nd, 4th, and 6th graders whose reading was analyzed were able to read two complete, difficult, unfamiliar stories each, and afterwards to discuss those stories. Several factors seem to affect the varying proficiency with which they read and talked about the stories: the ESL proficiency of the readers and the syntactic and semantic difficulty of the text appear to be the primary determiners of the subjects' comprehension of the materials.

The Miscue-ESL Project was a lengthy, in-depth study of the ESL reading of four language groups. The project began in 1973 at the Reading Miscue Research Center, Wayne State University, under the direction of Kenneth S. Goodman; this paper is a preliminary report of some of the major findings.[1] The Miscue-ESL Project was prompted by the results of an earlier study (Goodman and Burke, 1973) in which the miscues of children from 2nd, 4th, 6th, 8th, and 10th grades were analyzed.

One of the major points of that report was that the reading process was the same for all subjects, regardless of race, age, or reading proficiency. That is, all the subjects clearly used three cuing systems-graphophonic, syntactic, semantic—and clearly followed the basic process of sampling from these systems, predicting, and confirming. There was wide discrepancy among individuals as to how effectively and how efficiently this process was carried out, but the process itself did not differ from individual to individual. The result prompted the question, "Are there universals in the reading process?"

As a first step, the question was restricted to English, and rephrased as "Are there universals in the reading process when the reading is in English?" Four groups of subjects, children in 2nd, 4th, and 6th grades whose first languages were not English, were selected. These groups spoke, respectively, Arabic, Navajo, Samoan, and Spanish as their first language. The Arabic speakers were recent immigrants, primarily from Lebanon, to an urban suburb of Detroit. The area in which many of the Arab immigrants live is under the shadow (and the smoke) of one of the world's largest industrial plants. About half of the children at the school from which our subjects were drawn spoke Arabic, according to the school principal. At the time the data were collected, only one teacher spoke Arabic; he taught a small ESL class. Many of the stores and restaurants in the area offer Arabic specialities, and many of the signs outside and within these places are written in Arabic. This area is somewhat bilingual in the sense that two languages are used by many of the residents daily.

---

[1] The final report will be available early in 1978 from Dr. Goodman at the University of Arizona, Tucson, Arizona.

The Navajo-speaking subjects were residents of a boarding school run by the Bureau of Indian Affairs in the Arizona desert. None of the teachers spoke Navajo, although some of the teacher-aides did. The school was physically isolated from both Navajo-and English-speaking communities. There was no bilingual program, and the only ESL program was a cursory one in 1st grade. The children were members of two mono-lingual communities--Navajo in the summer, English in the school year.

The Samoan-speaking subjects were recent immigrants from American Samoa to an urban, polyglot community—Honolulu, Hawaii. They were acquiring a dialect of ESL, Hawaiian-Creole, where the data were collected. They lived in a housing project close to their school; the school had no bilingual program and a very small ESL program.

The Spanish-speakers were born in the small ranching community (population about 2,000) in east Texas where the data were collected. About a third of the school and the community are native Spanish-speakers. There are a few Spanish signs on grocery items, menus, and store windows. The community is partly bilingual in the sense that two languages are commonly used by many of the people, though with varying degrees of fluency. The school these subjects attended had no ESL classes, and at the time of data collection, the school was just beginning a bilingual program in the first grade. None of the subjects in this study had had a bilingual program; few had had formal ESL training.

With the exception of some older Arabic-speakers, who had attended school in their native land, and were literate in their language, all of the subjects in this study had received all reading instruction in English. English was the sole medium of instruction in all the selected schools.

The four groups were chosen for their diversity. They were all ESL-speakers, but they were drawn from urban and rural communities, bilingual and monolingual communities, native American and immigrant groups, Indo-European and non-Indo-European language backgrounds, from groups who had a long history of written literate and those who had none. If the study showed that the reading process was basically the same for each group, despite these wide differences, then the suggestion that there are universals in reading, at least in English, would be supported.

Ten students from each grade, 2nd, 4th, and 6th from each group were chosen, but not through statistical sampling. At each school, teachers were asked to choose the "most average" readers from their classes, that is, the students they considered the most average for their class in that school. It became apparent immediately that the teachers were selecting their best readers, and so the instructions to the teachers were changed. They were asked to list their ten best readers and their ten worst readers; ten subjects were then drawn from those not listed in either group. Each subject read two stories, one on one day, the second a day or two later. (A list of stories each group read is included as Appendix A.) The students were instructed to read the story without any assistance, and when they had finished the story, to tell the researcher everything they could remember about the story. The reading, done orally, and the retelling were both tape-recorded. All data were collected by researchers trained by Dr. Goodman at the Reading Miscue Research Center, and all data were sent there for analysis. Four subjects from each sub-group of ten were selected for analysis, primarily on the basis of the clarity of the tape-recording, to a lesser extent on the basis of other factors, such as gender or further information about the subject. For example, one subject who could not be used had been characterized by his teacher as a recent immigrant from Lebanon who spoke Arabic at home, but actually the boy had been born in Modesto, California, of a Brazilian mother and Lebanese father, and they spoke Portugese at home.

The reading and retelling of these 48 children, on two stories each, was analyzed using the Goodman Taxonomy. The Taxonomy is described and explained in detail in Allen and Watson's *Findings of Research in Miscue Analysis* (1976) and will not be explained at length here. Basically, the Taxonomy asks fifteen questions of each deviation from the text, each miscue. The answers to these questions are coded in Fortran, and then subjected to various computer programs.

I think the most important thing this study shows is that these children are better readers than they are usually given credit for. First, they are generally accurate readers: on an average they made miscues on less than 20% of the text, reading the other 80% just as it was printed. The measure of quantity of miscues that the Goodman Taxonomy uses is Miscues Per Hundred Words (MPHW). MPHW is arrived at by taking the number of different deviations from the text, dividing by the number of words in the text, and multiplying by 100. Identical repeated miscues, such as the repeated substitution of *Keeko* for *Keoki* are counted as one miscue. The range of MPHW for the ESL subjects was seven to twenty; their mean MPHW was about ten. The average, then, of what was read accurately, was 90% for all three grades, all four groups.

Second, when the children do deviate from the text, they often produce miscues which make sense in that text. They substitute names, for example, calling Tupa the shark *Tabu*, calling Mr. Barnaby *Mr. Barnberry*. Sometimes they read contractions as two words, saying *I am* for *I'm*. The Goodman Taxonomy codes these sorts of miscues, i.e. miscues which make sense in the text, as semantically acceptable. Table 1 shows the means and ranges of percent of non-dialect miscues that were coded semantically acceptable for each group: this includes all three grades and both stories.

Clearly, there is a great deal more variation within each group than between

TABLE 1

Percent of Semantically Acceptable Miscues by Groups.

Group	Mean	Range
Arabic	36%	18%-53%
Navajo	41%	20%-48%
Samoan	36%	16%-66%
Spanish	41%	20%-69%

groups, but the means suggest that, generally, at least a third of the miscues, before correction, make sense.

Third, when their miscues do not make sense, these children often regress and try again. When a child regresses and succeeds in producing the text item(s) accurately, the Goodman Taxonomy codes regression as *correction.* The measure of quality used by the Taxonomy is a percentage of the miscues which were semantically acceptable or were corrected. This figure is called the Comprehending Score. The next two examples would both be included in this score. The first is semantically acceptable:

Reader: Henry did not have a pet.
Text: Henry didn't have a pet.

The next example is semantically unacceptable, but is corrected, so it too would be included in the comprehending score:

©
Reader: didn't not
Text: Henry didn't have a pet. (Scott-Foresman, 1971, *Henry's Choice*, p. 1)

The mean Comprehending Scores for all grades on both stories are: Spanish—60%, Navajo—51%, Arabic—49%, and Samoan—47%. On the average, then, about half of these children's miscues either made sense or were corrected. Again, differences between individuals within groups were greater than differences between groups.

The fourth indicator of these ESL-speakers' ability to read unfamiliar and difficult text is their percentage of syntactically acceptable miscues. The average for each group is 55% or higher, and this is before correction. The range within groups and within grades is high—30% to 66% for Navajo second graders, for example: this is another indication of the greater variation within groups than between them.

The early results of this study, then, do not distinguish any one of the language groups involved as being better or worse readers of ESL. These results do show that on the average, these ESL-speakers read accurately; they produce meaningful miscues; they can correct when their miscues lose meaning; they are handling English syntax. The ESL-speakers in this study do not differ greatly from the native English speakers in the Detroit study (Goodman and Burke, 1973) in any of these aspects, except that their Comprehending Scores tend to be slightly lower. The number and types of non-dialect miscues the ESL-readers make are much more similar to those made by subjects in the Detroit study than they are different. To the question, "Are there universals in the reading process when the language is English," we can say, "It looks like there are." The focus of this paper, however, is not on what this study tells us about the reading process, but on what it tells us about the reading of ESL-speakers.

All the children made miscues which were obviously due to the fact that Standard English was not their first language or dialect. These miscues were phonological, grammatical, and lexical. The phonological miscues were not coded. That is, the pronunciation of *teacher* as *teasher* by a Spanish-speaker was noted on the transcript of the story, but was not analyzed. There are two reasons for this. First, control over the phonological system of English generally seems to follow control over the syntactic system, for both native and ESL-speakers, so that, although we expected the 6th grade Arabic speakers, for example, to read this sentence in the passive, we didn't expect them to distinguish /p/ from /b/.

Reader: At last the carbet was finished.
Text: At last the carpet was finished. (*Fareedah's Carpet*, p. 12)

Second, coding phonological deviations from an idealized Standard English pronunciation would not have told us anything about their actual reading. How much can one learn about a person's reading by noting how he pronounces *tomato* or *rodeo*?

Grammatical and lexical miscues caused by English being their second language were coded in the category called *dialect*. All dialect miscues are coded as grammatically and semantically acceptable, because the Taxonomy defines "acceptability" as acceptable within the reader's dialect. Most of the grammatical dialect miscues these children made are just what ESL teachers would expect: inflectional suffixes such as past tense *-ed*, plural *-s*, possessive *-'s* were not pronounced. A few of the children indicated that they had been instructed to "sound out the endings," because they produced supercorrect dialect miscues: *walkeded*, *slumpeded*, *lookeded*. Since some children produced more dialect miscues than others, and since we wanted to code an equal number of miscues for each subject, the dialect miscues were not counted in arriving at the total of the first fifty miscues. That is, each child had his/her first fifty non-dialect miscues coded; with dialect miscues included, one child had 52 total miscues coded, another 64.

The number of dialect miscues (ESL-miscues) varied widely from child to child. The Spanish-speaking group had a slight tendency to produce fewer dialect miscues than the other three: the average number of their dialect miscues per hundred words was one; the other groups' averages ranged from one and a half to two. These differences between groups are not significant. The types of ESL-miscues were similar for each subject and for each group: the large majority of these miscues (about 80%) were the substitution of a null-form for an inflectional suffix; some were the substitution of the base form of a verb for an irregular past or participle.

Taking means across grades and across stories lumps together ages and stories, and perhaps obliterates important differences. A clearer comparison can be made by looking at one grade reading the same story. All the 4th graders read "Freddie Miller, Scientist". Table 2 lists three measures of quantity and quality for this grade on this story.

TABLE 2

Measures of Quality and Quantity of Miscues on "Freddie Miller".

Group	MPHW	Comprehending Score	ResMPHW
Spanish	7.3	58%	2.48
Arabic	10.6	52%	4.7
Samoan	12.2	44%	5.78
Navajo	18.9	40%	9.06

The Comprehending Score, the Taxonomy's measure of quality, does not indicate quantity of miscues, as MPHW does. A measure that combines quality and quantity is the Residual Miscues Per Hundred Words (ResMPHW); this is the number of semantically unacceptable miscues which were not corrected per hundred scores. It is inversely related to the Comprehending Score. An ideal Comprehending Score is 100%; an ideal ResMPHW is 0.0. What could account for the very wide differences between the Spanish and Navajo readers on this story? The Spanish-speakers do not have much higher percentages of syntactic or semantic acceptability, indicating that their miscues are not better quality than those of the Navajos; that is, both groups are making the same *types* of miscues.

TABLE 3

Means and Ranges of Syntactic and Semantic Acceptability.

	Navajo	Spanish
mean syntactic accept.	58%	57%
range	44%-72%	49%-67%
mean semantic accept.	35%	35%
range	22%-42%	28%-39%

A comparison of their correction percentages strongly suggests that at least part of the answer lies in their ability and willingness to regress and correct. The Navajos corrected only 7% of their miscues, and only 2% of the semantically unacceptable miscues were corrected; the Spanish-speakers corrected 30% of their miscues. Moreover, the Navajo 4th graders successfully corrected only 6% of their partial attempts; the Spanish 4th graders 25% of theirs. A partial attempt is less than a full word; the correction of a partial suggests that the reader has recognized that s/he was miscuing and was able to correct that miscue before it was completed.

Ⓒ
Reader: Poor Free-
Text: Poor Freddie was in trouble again.

The much higher percentage of correction by the Spanish-speakers, of both full and partial miscues, together with the much lower number of MPHW made by the Spanish-speakers accounts for the large difference between their ResMPHW and the Navajos'. The Navajos miscued on almost one word in every five; the Texans made less than one-half that many MPHW. Of the Navajos' miscues, almost one in every ten words lost meaning and was not corrected. This is much higher than the Spanish-speakers' miscues; the number of uncorrected, semantically unacceptable miscues for them was a little over one in every 50 words. That is, although both groups made the same type of miscues, the Navajos made many more, and corrected many less than the Spanish-speakers.

These figures, MPHW, correction percentages, and so on, do not identify the underlying causes of the disparity between these two groups. These underlying reasons perhaps have something to do with the story itself; perhaps with the general ESL proficiency of the readers; perhaps with the status and general use (or non-use) of the students' first language in their schools and communities.

"Freddie Miller, Scientist" is about a young boy who continually gets into trouble because of his mechanical and chemical experiments, but who finally makes his family proud of him by using his "scientific" knowledge to make a flashlight for his little sister, who is locked in a dark closet. It seems to me that the Navajo children at the boarding school are less likely than the Texas ranch children to be familiar with the idea of a boy having a chemistry set, his own workbench in the cellar, or his using a kitchen stepladder to drop a flashlight through a transom into a closet. Do hogans have transoms or kitchen stepladders or cellars or chemistry sets? Does the boarding school? I suggest that, for Navajo reservation children who have been at boarding school for four years and have not lived in town, reading this story about a "typical" American family from an old standard basal text is in many ways like our reading Chaucer's "Prioress's Tale." What the people in the story say and do, their whole way of life, is foreign. This foreigness makes it very hard to predict what will happen in the next paragraph or page. Prediction is also hindered in some spots by the syntax. This passage, for example, occurs when Freddie's sister is stuck in the hall closet and Freddie is in the cellar.

> **. . ., he heard his sister's voice calling,**
> **"Freddie! Freddie!"**

> "Where are you?" he shouted.
> "In the hall closet!" came Elizabeth's tearful reply.
>
> . . .
>
> His sister's cries grew louder. "Don't leave me alone.
> It's dark in here."

The syntax of "came Elizabeth's tearful reply" forces readers to miscue. One predicts that "in the hall closet" will be followed by a dialog carrier like "said______" or "______said," but there is no real dialog carrier in this sentence. Instead we have a clause with inverted order (verb-subject); the verb is an intransitive one; the heavily modified subject (reply) is not even an animate noun, much less a person's name. In addition, "tearful" usually modifies animate nouns; here it is metaphorically used to modify "reply." It's easy to see why so many readers, native English speakers as well as ESL-speakers, produce: "In the hall closet," s-, came Elizabeth (pause) tearfully replied."

"His sister's cries grew louder" is another stumbling block for most readers. It too is in the position of a dialog carrier, but is not one. One predicts, "His sister said" or "His sister cried" or even "His sister cries," and the graphophonic information certainly supports the last two predictions. Some readers produce "His sister cried louder;" others make a couple of attempts and then seem to give up: "His sister cried, cries, sister cried lo-, grew louder." One must be a very proficient speaker of ESL indeed to get to deep structure successfully through such syntactic mazes.

In addition to the unpredictability of the general setting and plot and of some syntactic structures, there is the unpredictability of lexical items. When reading teachers say a child "knows" a word, they usually mean either the child understands the concept behind the word, or the child says the word accurately when s/he sees it, or both. But in "Freddie Miller," as in most stories, one word may have more than one meaning. The word *allowance*, for example, occurs three times in "Freddie." Mrs. Miller tells Freddie, "I want you to save half your allowance;" three lines later the narrator says, "After the cut in his allowance, Freddie's chemistry experiments narrowed to those safely outlined in a library book." (There's another syntactic maze.) And on the last page of the story, Mrs. Miller tells her husband, "We must make some allowance for experiments that do not turn out so well." I wonder how many of the Navajo 4th graders know the concept of allowance as weekly pocket money; I wonder if any know what "make allowance for" means. It's true that lexical items are generally unpredictable, but when this unpredictability is accompanied by the reader's inability to translate the item into something meaningful, then the reader's prediction will be a very tentative guess, and s/he cannot test to confirm that guess. If readers don't know the meanings a word has in different contexts and if they don't get enough context to enable them to deduce the meanings, they cannot test by asking, "Does that make sense?" The readers are then forced to reply on very surface information, or graphophonic information, which slows down the reading and makes prediction even harder.

In order to read with any proficiency, one must be able to predict, to make a reasonable guess as to what's coming next. The basis for prediction is partly what's in the text, but much more what's in the head of the reader, what s/he knows about the language and the world that the author is presenting. When the reader knows little about the world being presented, and when the author's language differs greatly from the reader's, the reader is going to have a difficult time and probably will make many miscues, and be unable to correct them. The reader cannot correct unless s/he recognizes that a miscue doesn't make sense; if *nothing* s/he's reading makes sense, there's no way of knowing which oral responses to the text are OK, and which aren't. While the world of "Freddie Miller" is not the world of east Texas Spanish-speaking ranch children, it is even less, I think, the world of reservation-boarding school Navajos. The language of "Freddie", although admittedly difficult in spots for any reader, is generally closer to the English controlled by the Spanish-speakers than to the English of the Navajos. Evidence for this comes mostly from the retellings: the Texas 4th graders volunteered lengthy paragraphs when asked to tell everything they could remember about the story; the Arizona 4th graders volunteered one or two sentences, sometimes only one or two words.

The Spanish-speakers were told before reading that they would retell the story in Spanish and in English; the Navajos only in English. What effect might this open acceptance of Spanish have had on the student's willingness to talk? The Spanish-speakers, like the Arabic speakers, live in a community where many of the people they know use two languages daily; the Navajo-speakers come from homes where Navajo is used almost exclusively, and live in a school where all the teachers, all Anglos, speak English exclusively. These differences in the use of two languages indicate two different attitudes towards the children's language, their culture, and themselves. What effect does this have on the children? On their reading? This study cannot answer those questions, because it did not investigate them. I hope others will.

What this study does answer, or at least begins to answer, are these questions: Can ESL readers, many of whom have not really mastered English yet, read with comprehension? Yes. Is ESL reading proficiency determined by one's first language? No. Are some aspects of the reading process universal? Tentatively, yes.

The Miscue-ESL project has implications for both research and teaching. More research into the reading process and into second language involvement in the reading process is called for. The results of this study suggest that there are universals in reading English; are there universals in reading any language? Sarah Lopez's research (1977) with Spanish-speaking subjects reading in Spanish suggested that the reading process in Spanish does not differ from the process in English. Studies need to be carried out in languages which do not use the Roman alphabet, such as Arabic, and in non-alphabetic writing systems, such as Chinese. Studies focusing on second language effect on reading, similar to the Miscue-ESL project, need to be carried out on different language populations. This research was a descriptive, in-depth study; others using statistical sampling techniques need to be done, for, although the results of this study suggest interesting insights, the results cannot be generalized. With growing interest in bilingual education and increasing fundings, bilingual educators and researchers could analyze the first- and second-language reading of a sample of students in a school committed to a bilingual program, and compare that to the first- and second-language reading of a sample in a school which has no bilingual program.

One of the major implications this study has for instruction is not in the results, it seems to me, but in the collection of data. One hundred twenty children were asked to read two complete stories each without any help and to talk about what they had read, and they did. A hundred and twenty children, 2nd, 4th, and 6th graders, from widely different language groups and cultural groups—every one of them read two complete, unfamiliar stories without help, and read them well enough to be able to talk about them. Even the poorest reader, one whose retelling score was very low and whose uncorrected miscues often destroyed both meaning and syntax, was able to finish his stories and to relate the major events. A few of the children said that this was the first time in their lives they had ever read a story all the way through at one sitting; they'd never been allowed to before, because someone was right there, stopping their reading whenever they hesitated or deviated from the text. The implication is clear, I think: our ESL students can read if we'll just let them. I remember the reaction I got from a TESOL Workshop when I showed the first page of a story as it was read by one of the Arabic-speaking subjects in this study: the Workshop participants were almost unanimous in condemning the story as too hard, and they were aghast when I told them I had allowed the student to finish the story, all fourteen pages. Most of them, they said, would have taken away the hard story and replaced it with something easier. But the student did read the whole story, and her retelling was excellent (Rigg, 1977). We do our students no service when we interrupt their reading in order to correct them. We are not helping them become good readers by treating reading as an exact process requiring three people—author, reader, and teacher. It is easy to confuse reading with reading instruction, but the two are separate processes and should be treated so. Reading is what the student does alone, with the text. Reading instruction is what the teacher does with the students to help them when they read. Many teachers feel they aren't earning their pay if they're not talking, but actually, it's harder work to remain silent when a student hesitates or makes an error than it is to jump in and read it for him. Allowing the reader to read without interfering strongly implies that s/he *can* read, and that the teacher recognizes that. It also lets the reader develop and use the strategies of prediction, confirmation, and correction, strategies which can be developed only through reading. In this respect, reading is like riding a bicycle; no amount of formal instruction can teach one as much as does that first wavering solo down the driveway.

A second aspect of the data collection that can be used in the classroom is the retelling. The students in this study were told before they began reading that they were expected to relate everything they could about the story after they had finished reading it. This instruction established comprehension as the purpose of the reading. It was tantamount to saying, "You'll have to think about what you're reading, because you'll be explaining it to me." The reader who knows s/he will retell must strive to understand the material well enough so it can be integrated into a coherent whole for presentation later. Just pronouncing the words won't do. Instructors who believe that good reading is a marvelous public performance, with every word pronounced exactly right, and with "beautiful expression," will want to train the students to be good performers, but instructors who think good reading is comprehension of the material will want to train for comprehension, not for performance. The two—comprehension and performance—are not automatically and mutually exclusive, but emphasizing perfect performance can reduce comprehension. "When you finish reading that to yourself, tell someone what you read." That is the kind of instruction that trains for comprehension.

Few, if any, teachers have time to listen to every student retell everything they read, but, of course, the retelling doesn't have to be to the teacher; it can be to a classmate. Students who have read the same story can share and compare their interpretations; those who have read different material can recommend what they have enjoyed.

A third implication from the Miscue-ESL project for teachers is that they can use oral reading as a window on the reading process. Teachers can get a view of their students' strategies by listening to their students read unfamiliar texts and by asking themselves, "Do the student's miscues make sense in this story? Do the sentences sound like his/her English? Does the student self-correct everything, even those miscues which retain meaning?" I don't mean the teacher should call up a reading group and have each one read a sentence or two aloud; that doesn't tell the teacher very much and it's often horribly embarrassing for the students. I mean that once every week or two the teacher can sit quietly with a student for 10 minutes while s/he reads aloud (without help, remember). If the teacher can tape-record the reading and analyze each miscue later, perhaps using Y. Goodman and C. Burke's *Reading Miscue Inventory* (1972), so much the better. But even without equipment or extra time, it is possible to get a rough measure of the student's reading by counting all the sentences s/he read, and noting what percent made sense as the student read them. (Y. Goodman, personal communication).

Finally, I want to point out a basic strength of this study—its assumption that reading is a language-based process of communication. This research did not ask for recognition of isolated words, nor for "perfect pronunciation" (whatever that is) of letter combinations. It did not deal with hypothetical subjects or with contrived sentences demonstrating syllogistic reasoning. The Miscue-ESL project studied real children reading real stories. The methods of collecting and of analyzing the data did not focus on exactness or on sets of skills, but on communication and on comprehension. A vital assumption underlying this research is that reading should make sense. The implication for teachers and for researchers is clear.

REFERENCES

Aaron, I. *et al.* 1971. *Henry's choice.* Glenview, Illinois, Scott-Foresman.

Allen, D. P. and D. Watson. 1976. *Findings of research in miscue analysis.* Champaign Urbana, ERIC/NCTE.

Betts, E. A., and C. M. Welch. 1963. Freddie Miller, scientist. *Adventures here and there; book 5.* New York, American Book Company.

Goodman, K. S., and C. Burke. 1973. *Theoretically based studies of patterns of miscues in oral reading performance, final report.* Project No. 9-0375, Grant No. OEG-0-9-320375-4269, U.S. Office of Education.

Goodman, Y. and C. Burke. 1972. *Reading miscue inventory.* New York, Macmillan.

Larson, J. R. 1966. Fareedah's carpet. *Palace in Bagdad.* New York, Scribner's.

Lopez, Sarah. 1977. Children's use of contextual cues. *The Reading Teacher 30,* 7, 735–740.

Rigg, Pat. 1977. Reading in ESL. In J. Fanselow and R. Crymes (eds.), *On TESOL 76,* Washington, D.C., TESOL, 203–210.

APPENDIX A

Author	*Publication Details*	*Book*	*Story*	
Emmet A. Betts C.M. Welch	N.Y.: American Book Co., 1963	*Around Green Hills*	The Big Surprise by M. Gartler	36-40
		Up the Street and Down	Two New Hats by E. McWebb	134-38
		Beyond Treasure Valley	Kitten Jones by R.G. Plowhead	60-66
		Adventures Here & There (Book V-3)	Freddie Miller, Scientist by Betts and Welch	61-68
		Adventures Here & There (Book VI)	My Brother is a Genius by W.D. Hayes	246-56
Ira E. Aaron, et al.	Glenview, Il: Scott-Foresman, 1971	*Henry's Choice*	Henry's Choice	16 p.
Albert J. Harris, Sr., et al.	N.Y.: Macmillan, 1966	*The Magic Word* (Book IV)	Royal Race by R. Eskridge	356-67
Helen M. Robinson, et al.	Glenview, Il: Scott-Foresman, 1967	*Open Highways* (Book VII)	Ghost of the Lagoon by A. Sperry	395-413
	Boston: Houghton-Mifflin, 1968	*Images* (Grade 5)	And Now Miguel by J. Krumgold	94-110

APPENDIX A--*(Continued)*

Author	*Publication Details*	*Book*	*Story*
Scott O'Dell	N.Y.: Dell, 1973	*Sing Down the Moon*	Sing Down the Moon by S. O'Dell 137 pp.
Mary Perrine	Boston: Houghton-Mifflin, 1968	*Salt Boy*	Salt Boy by M. Perrine 31 pp.
Helen Rushmore	Champaign, Il: Garrard Publ., 1972	*Sancho, The Homesick Steer*	Sancho by H. Rushmore 63 pp.
	Lyons & Carnahan	*The Almost Ghost & Other Stories*	My Name is Miguel by Fay & Clifford 55-74
James Holding	N.Y.: Morrow, 1963	*Mr. Moonlight & Omar*	Mr. Moonlight & Omar by J. Holding unpaged
Jean R. Larson	N.Y.: Scribner's, 1966	*Palace in Bagdad: Seven Tales from Arabia*	Fareedah's Carpet by J. Larson 46-59

Lessons Learned from Reading a Vietnamese Text: The Importance of Pre-texts in Second Language Reading

John C. Schafer[1]

Reading and Learning Skills Center
University of Michigan

Until recently most linguists in the U.S. studied sentences isolated from texts and most literary critics (especially New Critics) analyzed texts without attempting to explain what properties the particular text being analyzed shared with other texts in the same genre, or with all other texts belonging to the Anglo-American tradition. The popularity of sentence-based linguistics has made it difficult for ESL reading teachers to acknowledge that, in order to read, one must know more than vocabulary and syntax; the popularity of the New Critics' model of text analysis has made it difficult for them to perceive the value of making students aware not only of a text's internal relations but also its external relations to other texts in the culture. A new model for the analysis of texts which emphasizes external as well as internal relations is briefly described. One aspect of this model, the emphasis on the relation of a text to pre-texts, is illustrated by an analysis of a Vietnamese text. The paper concludes with a discussion of the important pre-texts in American culture and with the suggestion that ESL teachers expose their students to these pre-texts so they can develop the text competence necessary for reading.

The disciplines of rhetorical, literary, and linguistic analysis are becoming more and more to look like one discipline. Though they remain, at least in America, administratively compartmentalized (or departmentalized—in Speech, English, and Linguistics), rhetoricians, literary critics, and linguists are now all engaged in the same activity: the analysis of texts. The time seems right for a multidisciplinary approach to text analysis that would draw on the work that has been done in different disciplines. Such an approach might impress researchers with the advantages of cooperation. Since Professor Alton Becker of the University of Michigan has worked out an approach to text analysis that is truly multidisciplinary and that avoids, I think, many of the problems of other more compartmentalized approaches, I would like to outline his approach very briefly and illustrate one of its four main emphases, its emphasis on pre-texts, by commenting on a Vietnamese text. The case to be argued here is that a knowledge of pre-texts is a crucial part of the text competence of a native speaker. Without this knowledge one can not read as Frank Smith (1971) and Kenneth Goodman (1973) say a fluent reader reads, i.e., by making predictions that are later verified or modified.

---

[1] I would like to thank Professor Alton Becker, Department of Linguistics, University of Michigan, whose ideas form the basis of the approach taken here, and Professor Joyce Zuck, English Language Institute, University of Michigan, who made many constructive suggestions.

But first some problems of definition. *Discourse analysis* and *text analysis* have been used more or less interchangeably to refer to attempts to identify and describe the structure of units of verbal behavior larger than the sentence. Discourse analysis has been the more common term in English speaking countries while text analysis has been used on the continent of Europe to refer to the same activity. Although recognizing that discourse analysis and text analysis have often been used interchangeably, Sandulescu points out that discourse analysis usually refers to "*data-centered*" approaches to the study of *spoken language* while text analysis commonly refers to "*model-oriented*" approaches to the study of *written language* (Sandulescu, 1976, p. 349-365). Widdowson suggests a slightly different terminological distinction (Widdowson, 1976, p. 57-58). He recommends that discourse analysis be used to refer to investigations of the communicative aspects of an instance of language and text analysis to refer to studies of the formal properties of a series of connected sentences.

Though Sandulescu's and Widdowson's distinctions may be useful, they are not utilized here because they would seem to lead to a compartmentalization of efforts, when, as mentioned above, I think a cooperative, multidisciplinary approach has more merit. I will therefore dispense with the term discourse and use text to refer to any instance of language, spoken or written, which has an internal structure and is perceived by the people of a culture to have a beginning, a middle, and an end. Text analysis is the study of such an instance of language, including both its communicative and formal aspects.

An Approach to Text Analysis

Becker's multidisciplinary approach starts with the assumption that meaning is not inherent in a thing itself but rather emerges from its relations to its context (Becker, 1977, p. 1-3). What kinds of relations are important? Becker says to fully understand a text one must understand four sets of relations:

1. The relation of parts of the text to each other: *co-text* relations. These are the relations examined by Halliday and Hasan who are interested in learning how cohesion is achieved in English texts (Halliday and Hasan, 1976).

2. The relation of the text to previous texts, to *pre-texts*.

3. The relation of the text to the *intention* of the speaker/writer with intention defined as "the relations of the creator to the content of the text, the medium, and to the hearers or readers" (Becker, 1977, p. 3). These relations have been examined by rhetoricians, by speech and text act theorists, and by literary critics interested in interpretation theory.

4. The relation of the text to the world, to non-literary things and events, to its *context*. Context is often used in a more general way to refer to 1 through 4 above; Becker restricts its meaning to this fourth set of relations, the relations that are commonly called reference by linguists.

It is the second set of relations, the relation of a text to pre-texts, that I would like to discuss in more detail. I would like to concentrate on this set of relations because it is a set that other text analyzers have under-emphasized. For example, because they were reacting against historical criticism, the New Critics had very little interest in this set of relations. Structuralist literary critics and modern linguists, because they are interested in synchronic not diachronic approaches, have also not paid much attention to pre-texts in their analyses. (For a counter example, see Culler, 1975, p. 131-160.) Realizing there were other relations to discover besides historical ones has been liberating for both literary critics and linguists, but since all texts speak the past as well as the present, a text's relations with former texts are part of its meaning, a part that should not be put aside in our enthusiasm for understanding synchronic relations. In this regard it is interesting to compare Labov's work much of which is a warning that structural systems of the present should not be considered apart from historical change (Labov, 1972). Jakobson's comments on the Russian Formalists are also instructive. They knew, he said, that "shifting and change [in the relations between texts] are not only historical statements (first there was A, and then A_1, arose in place of A) but that shift is also a directly experienced synchronic phenomenon, a relevant artistic value. The reader of a poem or the viewer of a painting has a vivid awareness of two orders: the traditional canon and the artistic novelty as a deviation from the canon" (Jakobson, 1935, p. 87).

Not only famous linguists but our own common sense as well affirms the importance of pre-texts. We know, as Becker points out, that any cowboy movie is in many ways much more about previous cowboy movies than about actual men who herded cattle in the West. In regard to a movie like *Blazing Saddles* the relation of text to pre-texts is purposely made obvious, but, as Shklovsky pointed out, "not only parody, but also in general any work of art is created as a parallel and a contradiction to some kind of model" (Shklovsky, 1919, p. 53).

Nor is the situation much different for non-poetic prose. To understand much of the mundane prose we hear or read we must know more than just vocabulary and syntax: we must know the scripts, the term Schank uses for more stereotypic pre-texts (Schank, 1975, p. 4). A script is "a predetermined stereotyped sequence of actions that define a well-known situation" (Schank, 1975, p. 4). Schank asks us to consider the sequence:

> John went to a restaurant. He found a table and ordered a hamburger. Later, he paid and left.

Schank says unless we know the restaurant script we can not easily connect finding tables and ordering. The whole script that involves getting a waitress' (in America it's usually a waitress not a waiter) attention, a dialogue with the waitress, the waitress bringing food, etc., has been evoked but only parts of the total script are found in the surface structure of the above passage. In understanding a text like the short one above that evokes a script, the script becomes part of the story even when it is not spelled out. "The answer to the question 'Who served John the hamburger?' seems obvious because our world knowledge, as embodied in scripts, answers it" (Schank, 1975, p. 10).

A Vietnamese Text

One way to be impressed with how important it is in reading to know more than vocabulary and syntax is to try to read a non-English text oneself. In trying to read a Vietnamese text recently I found all four sets of relations described above to be crucial, but I will comment here only on the role of pre-texts. The text I have been reading is a 2086 line narrative poem entitled *Lục Vân Tiên*, which was composed in 1865 by a blind Vietnamese poet named Nguyen Dinh Chieu. I started with the following lines which I later realized were a prologue:

Text in Quốc Ngữ (Romanized script)

(1) Trước đèn xem chuyện Tây-Minh,

(2) Gẫm cười hai chữ nhơn tình eó le.

(3) Hỡi ai lẳng lặng mà nghe,

(4) Giữ răn việc trước, lánh dè thân sau.

(5) Trai thời trung hiếu làm đầu,

(6) Gai thời tiết hạnh là câu trau mình.

Translation

(1) Before the light I look at the story of Tay Minh

(2) And muse over the two deceiving words, human feeling!

(3) Attention everybody! Be quiet and listen,

(4) Recollect past mistakes, avoid bad consequences later:

(5) Men take loyalty/filial piety as your rule,

(6) Women take chastity as the word to improve yourselves.

One of the dangers of analyzing foreign language texts is that one mistakes the stereotypic for novelty because everything is new when one first starts to read texts in another language. This was what happened to me when I read *Lục Vân Tiên*. Since it was only the second Vietnamese narrative poem I had ever looked at, when I read the above lines I accepted most of the information presented in them as the original creation of Nguyen Dinh Chieu. But later in talking to Vietnamese informants and in reading other Vietnamese narrative poems, I was surprised to learn that much of the material in these six lines was stereotypic. It is stereotypic, for example, to refer to older books as the source of information for the story one is to tell, as the narrator does in the first line; it is stereotypic to muse over the meaning of characters, as the narrator muses over the meaning of the two characters for human feeling in the second line (often in fact the same characters are mused about: the characters for humanity, or feeling, or fate, or talent); it is also stereotypic to state in the opening lines a general rule of human existence that the present story is to instantiate, as the narrator does in the above prologue. If, as Frank Smith and Kenneth Goodman have suggested, "the information that passes from the brain to the eye is more important in reading than the information that passes from the eye to the brain," and if it is this behind the eyeball information that helps the fluent reader make predictions that he later verifies or modifies, then this explains why I was at a disadvantage when I read this text (Smith 1971: 9; see also Goodman 1973). I could make no predictions because I had no experience with previous texts, experiences that would have instilled in me a certain set of expectancies. A lot of learning to read foreign texts involves learning a new set of expectancies.

How can we help students acquire a good set of expectancies? One thing that should be done is to be sure they get exposed to different types of texts so they can develop different expectancies for different texts. We can encourage students to see co-occurrence relations among texts in a particular genre; we can point out, for example, that most newspaper articles place the most important information first, that term papers usually have a thesis statement at the end of the introduction, that epic poems often begin *in medias res.* Many teachers undoubtedly already point out such things but I think many of us overemphasize the uniqueness of the particular text being studied and underemphasize what it shares with other texts in its genre and with all other texts in the Anglo-American tradition. Influenced perhaps by the *explication de text* approach of the New Critics, the approach to text analysis inflicted on many of us in college literature classes, we treat each reading assignment as a completely fresh creation. We painstakingly analyze a text to reveal its internal relations but say nothing about its external relations, its relations to other texts in its genre, to other texts in the culture. In treating the work as completely unique we give the student little guidance in developing a set of expectancies, a set of "advance organizers" (Ausubel 1968: 148-9) that reading researchers believe a good reader must have (Rickards 1976).

In my own struggle to read this Vietnamese text, I soon found knowing prior texts involved more than just knowing other texts belonging to the same genre. It involved in addition knowing texts in other genres and how these texts related to the text I was analyzing. To understand *Lục Vân Tiên,* for example, I found I had to know something about *ca dao*, or folk poetry, the pre-texts for most verbal art in Vietnam. *Ca dao* lines becomes lines in poems and lines in poems become *ca dao*, both become pre-texts for each other, in complicated ways only a native researcher could unravel. Compare, for example, the following *ca dao* with lines 5 and 6 of Nguyen Dinh Chieu's poem:

> When you enter into this reincarnation,
> Men, take loyalty/filial piety, one on each shoulder,
>   and fulfill both these duties completely;
> Women, be faithful to the rule of chastity;
> Day and night be careful to avoid any imperfection.

But I found there were still other pre-texts besides other narrative poems and folk poetry that I had to know before I could understand this prologue. I found I had to know something about the ancient books of China, the *Four Books* [*Tứ thư*] and the *Five Classics* [*Ngũ kinh*]. How are these ancient Chinese texts evoked? Often by a kind of shorthand that takes some getting used to. In the above prologue, for example, the words *trung hiếu* [loyalty-filial piety], line 5, and *tiết hạnh* [purity-behavior], line 6, evoke in the minds of most Vietnamese the whole Confucian ethical system, a system that is described in the *Four Books* and *Five Classics*. In saying men should be loyal and filially pious and that women should be pure, Nguyen Dinh Chieu is not saying that men and women should have only these virtues; he is recommending a return to the entire Confucian moral system as a way of restoring order in the society.

To explain more precisely how ancient Confucian texts are evoked, however, it is necessary to discuss a special kind of pre-text, but perhaps text is not the right word for what I wish to describe. What I am referring to are certain classes and hierarchies that mediate between old and new Vietnamese texts (See Table 1). Some important classes are: the three bonds [*tam cường*], the five virtues [*ngũ thường*], the three female submissions [*tam tòng*], and the four female virtues [*tứ đức*]. Some important hierarchies are the hierarchy of four occupations and the order of activities of a superior man. The source for most of these Vietnamese classes and hierarchies are the Chinese *Four Books* and *Five Classics*. For example, the Three Submissions (III in Table 1) are described in the Classic of Rituals; the order of activities of a superior man (VI in Table 1) comes from the *Great Learning* (one of the *Four Books*).

Once I knew about these classes and hierarchies I found I had a very important set of handles that I could use to grasp many Vietnamese texts. But it wasn't enough to know only that these classes and hierarchies existed; I also had to know what their members were, something about the collection of meanings associated with each member, and the shorthand used in texts to evoke them.

Learning the members was no problem. They are known by most Vietnamese who will tell you, for example, what the Five Virtues and the Three Female submissions are if you ask them. These classes and hierarchies are

TABLE 1

Sino-Vietnamese Classes and Hierarchies

I. Tam cương	The Three Bonds	
Sino-Vietnamese	*Vietnamese*	
phụ/tử	cha/con	father/son
quân/thần	vua/tôi	king/subject
phu/phụ	chồng/vợ	husband/wife
II. Ngũ thường	The Five Cardinal Virtues	
nhân	humanity	
nghĩa	righteousness	
lễ	propriety	
trí	knowledge	
tín	sincerity	
III. Tam tòng	The Three Submissions (for women)	
tòng phụ	first to her father	
tòng phụ	second to her husband	
tòng tử	third to her son	
IV. Tứ đức	The Four Virtues (for women)	
công	work	
dung	physical appearance	
ngôn	correct speech	
hanh	proper behavior	
V. Hierarchy of the Four Occupations		
first	sĩ	scholar
second	nông	farmer
third	công	craftsman
fourth	thương	merchant
VI. The Order of Activities of a Superior Man (quân tử)		
first	tu thân	improve oneself
second	tề gia	improve one's family
third	trị quốc	make the country prosperous
fourth	bình thiên hạ	make the world peaceful

brought up often in informal conversations as well as formal debate in Vietnam, not as descriptions of present behavior and beliefs so much but rather as devices for measuring how far present behavior and beliefs deviate from earlier standards. As pre-texts they continue to exert an influence on present lives. David Marr has explained how twentieth century books on the proper behavior of Vietnamese women all begin with a discussion of the Three Submissions and the Four Virtues (Marr 1976). What the writers of these books tried to specify was how much a modern Vietnamese woman could deviate in behavior from the traditional code summed up by the words *Tam tòng tứ đức* [Three Submissions Four Virtues].

More recently Ho Chi Minh attempted to re-contextualize some of these Confucian classes and hierarchies. In doing so he evoked the prologue to *Lục Vân Tiên* and the Confucian Three Bonds. In the prologue to *Lục Vân Tiên* quoted above the words[2] *trung hiếu* [loyalty-filial piety] evoke, as I've said, the whole Confucian moral system, but more specifically they evoke two of the Three Bonds: *vua-tôi* [king-subject] and *cha-con* [father-son]. In the Confucian system one was loyal [*trung*] to the king and pious [*hiếu*] toward one's father. In fact loyalty to king was viewed as an extension of the piety extended toward one's father. From the point of view of a Marxist revolutionary then, two of the Three Bonds, king-subject and father-son, are "feudalistic" and need either to be repudiated, as they were at least for a while in China, or somehow re-contextualized, which appears to be what is being attempted in Vietnam. Using the same words (*trung-hiếu*) as the Confucian scholar Nguyen Dinh Chieu, Ho Chi Minh used to tell cadres: "We must be loyal [*trung*] to the Party and pious [*hiếu*] toward the people" [*Phải trung với đảng, hiếu với dân*].

Learning the collection of meanings associated with each of the members of these classes and hierarchies is a formidable task precisely because these sets and their members have been re-contextualized so many times both in China and in Vietnam. A somewhat less formidable but still challenging task that faced me when I tried to read the poem *Lục Vân Tiên* was to learn the shorthand the author used to evoke these classes and hierarchies. Sometimes they are referred to very explicitly as the *Tam cương* [Three Bonds] or *Ngũ thường* [Five Virtues] much as we would refer to the Ten Commandments or the Seven Deadly Sins. I had no difficulty in understanding what was meant when these classes or hierarchies were referred to in this straightforward manner. But to become alert to other ways they could be evoked I found I had to accustom myself to the devices of reduction, parallelism, and binary opposition in Vietnamese. When writers wish to evoke both the *Tam cương* [Three Bonds] and *Ngũ thường* [Five Virtues], instead of saying *Tam cương Ngũ thường*, they may reduce this four word expression and directly oppose the Bonds to the Virtues in the compound word *Cương-thường* [Bonds-Virtues]. This compound word then stands for all the moral obligations, including the bonds and virtues, stressed by Confucian doctrine. When writers wish to speak of the entire ethical system as it applies to both men and women, they pair the virtue of loyalty [*trung*], which was the virtue especially but not exclusively for males, with the virtue of chastity *[trinh]*, the virtue reserved for females, in the compound word *trung-trinh* [loyalty-chastity].

---

[2] For a discussion of the difficulties of defining a word in Vietnamese, see Thompson, Lawrence 1963. "The Problem of the Word in Vietnamese." *Word*, 19, pp. 39-52.

Reduction, parallelism, and binary opposition in Vietnamese do not operate just among the Confucian bonds and virtues. There are countless examples in *Lục Vân Tiên* and in the Vietnamese vernacular of pairings being used to refer to non-Confucian categories or qualities. (Table-chair [*bàn-ghể*] is furniture; end-beginning [*chung-thủy*] means faithful in marriage; mountain-water [*non-nước*] refers to the country of Vietnam.) Nor are the operations of parallelism and binary opposition restricted only to the level of the compound word; they are important features of the Vietnamese language at all grammatical levels—at the level of the compound word as we have seen above, at the level of the four syllable idiomatic expression, and at the level of the sentence (Nguyen Dinh Hoa 1965). One gets a suggestion of both parallelism and binary opposition in lines 5 and 6 of Nguyen Dinh Chieu's prologue:

(5) Men take loyalty/filial piety as your rule,
(6) Women take chastity as the word to improve yourselves.

Binary opposition and parallelism are also very important devices at higher levels of Vietnamese text organization. In the story that follows this prologue the hero Lục Vân Tiên exemplifies loyalty and filial piety and the heroine Kiều Nguyệt Nga exemplifies chastity. Lục Vân Tiên interrupts his improving himself (first in the order of activities of a superior man), which he is attempting to do by studying for the mandarin exams, to attend to his mother who is dying; thus he demonstrates filial piety. Later he serves his king by helping him put down an uprising; by doing so, he demonstrates loyalty to his king. At the same time Kiều Nguyệt Nga, the accounts of whose adventures are paralleled to and interspersed with those of Lục Vân Tiên, demonstrates the virtue of chastity. Thinking that her lover Lục Vân Tiên is dead, she takes a vow of chastity, deciding to remain pure forever to the memory of the man she had promised to marry.

Sinologists have commented on the harmonious relationship between parallelism as the most important feature of the Chinese verbal style and the Chinese conception of the world as divided between the principles of *yin* and *yang*, *yin* being the female principle and *yang* the male principle, the combination and interaction of which is supposed to account for all things. But this speculation concerns China not Vietnam. In any event one does not need to turn to China for sources of Vietnamese parallelism and binary opposition. According to the Vietnamese myth of creation, the Vietnamese people are descendants of the dragon king Lạc Long and the fairy princess Âu Cơ, King Lạc Long's mother's home was the realm of water; Princess Âu Cơ was born and raised on a high mountain. From their union, the union of the opposites of water and mountain, came the land and people of Vietnam. This story is perhaps the pre-text par excellence of Vietnam and might be compared to the story of creation, and of Adam and Eve, in Genesis.

In summary, in order to understand this prologue to the Vietnamese poem *Lục Vân Tiên* I found I had to know, among other things, how this text related to other texts—to other verse novels such as *Nhị Độ Mai* and *Kim Vân Kiều*, to *ca dao* folk poetry texts, to the *Four Books* and *Five Classics* of China, and to certain classes and hierarchies such as the Three Bonds and the order of activities of a superior man.

Pre-texts in American Culture

What are the important pre-texts for texts in American English? Could they be discovered and classified and taught systematically to ESL students? I don't see why not. I've been working on a list that so far includes: the Bible—especially the story of creation, the story of Adam and Eve, the life of Christ; some Mother Goose rhymes; some fairy tales—maybe Cinderella, Snow White, and the Frog Prince; selections from the writings of Sigmund Freud; something on Puritanism—perhaps a biography of Cotton Mather; the Constitution and the Bill of Rights; an Horatio Alger story; Owen Wister's *The Virginian*; a baseball rule book; a first aid handbook and popular book of remedies for colds and other common ailments; Emily Post's *Etiquette*; sayings from *Poor Richard's Almanac*; and assorted idioms, clichés, and other frozen forms.

As for scripts, the more pedestrian everyday texts, many are already taught as dialogues or situation drills in ESL speaking classes. Scripts such as "At the Restaurant," "At the Doctor," and "Student Meets Faculty Advisor" are part of an ESL teacher's stock and trade. To make these dialogues or situation drills interesting (perhaps interesting for ourselves) we often introduce deviations and turn them into what Schank calls stories (Schank, 1975, p. 4). Students may have trouble, however, distinguishing between expected and unexpected behavior. There may be some advantage in teaching students first the scripts, to give them a feel for what is stereotypic, before introducing stories that contain deviations from the scripts. There may be some advantage in making explicit the distinction between a *script* situation drill and a *story* situation drill, especially for students studying English abroad who aren't getting exposed to stereotypic American situations outside the classroom.

Though one shouldn't expect to find the same type of pre-texts in different cultures (or in different epochs of the same cultural tradition) one still keeps looking for them. After becoming impressed with the importance of classes and hierarchies in the Sino-Vietnamese tradition, I began to wonder if there were similar pre-texts in the Anglo-American tradition. I came up with the Ten Commandments, the Four Humours, the Four Freedoms, the Seven Deadly Sins, the Great Chain of Being. While I knew from previous experience as an English major that a knowledge of these classes and hierarchies was crucial in piecing together an understanding of the works of English literature, they didn't seem to be essential pre-texts for much modern writing. Undoubtedly this is because the role *a priori* categories play as pre-texts in our society is rather small. Science has deplenished our supply of eternal verities. New Classes such as Freud's Id, Ego, and Superego come along but soon they suffer the same fate as the Four Humours. Ethical categories sanctioned by religion last longer but even the Ten Commandments seem to be evoked less and less these days. Perhaps the pre-texts comparable to the classes and hierarchies of China and Vietnam, and of an earlier England and America, are not *a priori* categories but heuristic procedures, problem solving techniques, the scientific method: research the problem, develop an hypothesis, test it, announce your result—which won't be a law but might become a useful way of making sense of things until something better comes along.

Conclusion

One may ask: Does one really have to know all these things, all four of Becker's sets of relations, all of these pre-texts and scripts, to read a text in a foreign language? Can't one read a text if one knows the writing system, a fair amount of vocabulary, and sentence syntax? I think one does need to know all these things. Vocabulary and syntax aren't enough. And I think people in ESL have known for some time they weren't enough but the popularity of the theories of linguists who separated form from function and synchronic from diachronic relations, who took isolated sentences as their corpus for study, who had a very narrow definition of competence, has, until recently, kept teachers from saying so forcefully.

Of course there are different levels of mastery; people have spoken of a cline of bi-lingualism. Undoubtedly there are different degrees of understanding of a text. As teachers of reading, however, I think it may be dangerous to underestimate the number of things one needs to know to read texts in a foreign language. If we minimize the amount of knowledge necessary for reading we may move our students to a pre-mature closure. They may be encouraged to think they know enough when they've only scratched the surface. An analogy could be drawn to the ESL speaking teacher who encourages students to think they have mastered a sentence when, given a cue from the teacher, they can mouth it correctly in class. Instead of allowing students to reach closure at a low level of mastery, we reading teachers should encourage students to develop their "behind the eyeball" knowledge (Smith, 1971, p. 68-79), to begin to explore some of these other areas of knowledge that Becker describes. The hopeful thing about the coalescing of the disciplines of rhetoric, literary criticism, and linguistics around an interest in texts, as I see it, is that it may lead to an increase in multidisciplinary approaches to text analysis, approaches that should help make us aware of the competence we as native speakers possess that enables us to take meaning from English texts. This competence, which might be called text competence, is certainly broader than Chomsky's grammatical competence and broader also than what many people have in mind when they use the phrase communicative competence. With our consciousness raised regarding this text competence, we may be able to introduce to students, in more systematic fashion, what they need to know to read.

REFERENCES

Ausubel, David. 1968. *Educational psychology: a cognitive view*. New York, Holt, Rinehart, and Winston.

Becker, A. L. 1977. Text-building, epistemology, and aesthetics in Javanese shadow theatre. (To be published.)

Culler, Jonathan. 1975. *Structuralist poetics*. Ithaca, Cornell University Press.

Goodman, Kenneth. 1973. *Miscue analysis*. Urbana, NCTE.

Halliday, M. A. K., and Ruqaiya Hasan. 1976. *Cohesion in English*. Hong Kong, Longmans.

Jakobson, Roman. 1935. The dominant. In Ludislav Matejka & Krystyna Pomorska (eds.), *Readings in Russian poetics*, Cambridge, M.I.T. Press, 1971, pp. 79–81.

Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia, University of Pennsylvania Press.

Marr, David. 1976. The 1920's women's rights debates in Vietnam. *Journal of Asian Studies 35*, 3, 371–389.

Nguyen Dinh Hoa. 1965. Parallel constructions in Vietnamese. *Lingua 15*, 125–139.

Rickards, J. P. 1976. Processing effects of advanced organizers interspersed in texts. *Reading Research Quarterly 11*, 599–622.

Sandulescu, Constantin-George. 1976. Theory and practice in analysing Discourse. In Gerhard Nickel (ed.), *Proceedings of the Fourth International Congress of Applied Linguistics*, Vol. 1, Stuttgart, HochschulVerlag, pp. 349-365.

Schank, Roger C. 1975. SAM—A story understander. Yale A. I. Project, Research report #43, 1–42.

Shklovsky, Viktor. 1919. On the connection between devices of *syuzhet* construction and general stylistic devices. In Stephen Bann and John E. Bowlt, (eds.), *Russian Formalism*, New York, Harper and Row, 1973. Pp. 48–72.

Smith, Frank. 1971. *Understanding reading*. New York, Holt, Rinehart and Winston.

Thompson, Lawrence. 1963. The problem of the word in Vietnamese. *Word* 19, 39–52.

Widdowson, H. G. 1976. Discourse analysis, interpretative procedures and communicative language teaching. *Bulletin CILA* [Commission Interuniversitaire Suisse De Linguistique Appliquee], 23, Neuchatel, 57–63.

Shades of Meaning: Syntactic and Semantic Parameters of Cloze Test Reponses

Mark A. Clarke and Linda Burdell[1]

English Language Institute
University of Michigan

Cloze tests have gained wide acceptance in L2 research and teaching. Shown to correlate highly with standardized tests of ESL proficiency, they have been used to elicit L2 research data and to improve students' use of context in vocabulary and structure work. They have also proved useful as testing devices. Typically, tests are constructed by deleting every n*th* word from a passage; subjects are required to fill the resulting blanks with words which satisfy the syntactic and semantic constraints of the context. The purpose of this paper is to describe an analysis tool which specifies the linguistic criteria for judging the acceptability of cloze test responses. The tool contains three categories: syntactic acceptability (SYNAC), semantic acceptability (SEMAC), and semantic change (SEMCH). The code values for each category are defined and examples are discussed. This analysis evaluates the quality of errors produced by subjects, and provides insights into the processes used to produce responses. Implications for L2 research and teaching are discussed.

Originally developed to measure the readibility of prose (Taylor, 1953), cloze tests have gained wide acceptance in L2 research and teaching. Tests are typically constructed by deleting every n*th* word from a passage and subjects are required to fill each of the resulting blanks with an appropriate word. Cloze tests have been shown to be valid and reliable measures of ESL proficiency (Oller and Conrad, 1971; Irvine, et al., 1974; Stubbs and Tucker, 1974) and they have proved valuable in L2 teaching (Eskey, 1973, Plaister, 1973; Berkoff, 1976).

The optimum scoring method remains a matter of debate. In scoring native speaker performance the most efficient method is to count the number of times a subject produces the exact word used by the author (Taylor, 1953; Rankin, 1957). In L2 research, however, the tendency has been to use scoring systems which give credit to contextually acceptable responses. Numerous studies have shown very high, significant correlations between exact and acceptable scoring. Although research indicates that the time and effort required to produce an "acceptable" key do not yield significantly different results, researchers regu-

---

[1] We would like to thank Brad Arthur, Ann Borkin, and Jack Upshur of the English Language Institute, University of Michigan, for comments on an early draft of the analysis tool presented in this paper.

larly report both exact and acceptable scores. (In particular, see Oller, 1972 and Stubbs and Tucker, 1974 for discussions of the merits of exact/acceptable scoring procedures.) There are a number of reasons for reporting both scores. As Oller (1972) points out, requiring exact-word replacement not only makes test extremely difficult for non-native speakers, but it also requires insights which perhaps cannot be considered language skills. In addition, it seems intuitively justifiable to give more credit to an individual whose answers satisfy some of the constraints on a blank, than to one who has ignored all contextual requirements.

In spite of apparent consensus concerning the use of the acceptable scoring method in L2 research, little attention has been given to the issue of what linguistic criteria are to be used in assessing acceptability. In most cases, the authors merely state that tests were scored for exact and acceptable responses, with little or no mention of the procedure used to arrive at decisions of acceptability. It can generally be assumed that native speaker judgments were used, and that ungrammatical responses were rejected, as were responses which violated the instructions (ie: blanks, more than one word, native language responses). These assumptions constitute valid criteria for scoring cloze tests if the goal is to assess L2 proficiency or if the researcher is a non-native speaker (Stubbs and Tucker, 1974). For judgments on comprehension, or for item analyses which reveal varying levels of L2 competence, a more sensitive instrument is required.

A few researchers have attempted to develop objective procedures for evaluating cloze test responses. Darnell's (1970) "Clozentropy" procedure is a complicated statistical approach which allows one to compare a subject's performance against that of some target population. The method is of little value to the individual researcher or teacher who cannot test large numbers of native speakers to obtain the baseline data, or who does not have access to computer assistance. Two other scoring attempts, however, merit discussion.

Bowen (1969, p. 74) developed a five point scale for evaluating the cloze test responses of Amharic-speaking Ethiopian students. Each blank was allowed a maximum of four points:

4: original word restored or fully meaningful synonym or replacement
3: meaningful replacement, but idiom or grammar slightly strained
2: suggestive replacement, but idiom strained or minor grammar error
1: understandable answer, but unidiomatic replacement or strained idiom or minor grammar point
0: unacceptable answer, or no answer given

It is easy to see that judgments of "strained grammar" or "meaningful" versus "suggestive" replacement could vary greatly from investigator to investigator. In fact, Bowen (1969, p. 75) reports that a great deal of scorer-consultation occurred during the evaluation of the tests. The necessity of communication between scorers reduces the objectivity of the evaluation and subsequently, the replicability of the research by different investigators.

Oller has developed a differential analysis which to a large extent remedies the weaknesses of the above procedure (Oller, 1972; Oller et al., 1972). The system attempts to assess the degree of correctness of each response. Five subjective categories are used:

1: entirely acceptable; something a native speaker is very likely to say
2: contextually acceptable; violates no obligatory or demonstrable restrictions but constitutes unusual word order
3: violates no obligatory restrictions in the immediate phrase structure context but does violate some broader constraint(s)
4: violates obligatory selectional constraints in the immediate phrase structure context but in some way fits the larger context or could be changed slightly so as to become fully acceptable
5: violates obligatory constraints and seems unmotivated by any substantial degree of comprehension
6: blank (Oller et al., 1972, p. 9)

This scheme allows responses to be categorized as Correct (#'s 1&2) Partly Correct (#'s 3&4) and Incorrect (#'s 5&6). This procedure is more explicit in delineating the parameters of each category. However, the absence of specific coding instructions describing and/or giving examples of "unusual word order" or "selectional restraints" limits the usefulness of the tool for other researchers; the native speaker must establish his own criteria, and the non-native speaker is provided with no guidelines. In addition, the results produced by this procedure remain primarily quantitative. Its value as a scoring tool has been demonstrated (Oller, 1972, p. 153-157; Oller et al., 1972, p. 9-11), but if one wants to evaluate subjects' response strategies, or to diagnose learner difficulties for instructional purposes, finer distinctions are required.

Our purpose here is to describe in detail a procedure we have developed for the analysis of L2 cloze test responses. We hope that the paper will become a "code sheet" for cloze analysis, a practical working document that other L2 researchers can use in their work. For this reason we provide detailed discussions of coding instructions and examples. Common coding problems are mentioned and implications of the instrument for L2 research and teaching are discussed.

Syntactic and Semantic Parameters

As part of a larger study of L1/L2 reading behavior (Clarke, 1977) we have been involved in an analysis of the reading strategies of adult ESL students as revealed by their cloze test performance. For our purposes, "reading strategy" is operationally defined as a subject's use of context in responding to a mutilated text. Reading is viewed as "psycholinguistic guessing game" (Goodman, 1970) in which readers sample graphic, syntactic and semantic cues to confirm or reject their hypotheses about the meaning of prose. The assumption is made that cloze tests are valid measures of reading ability (Oller and Conrad, 1971; Irvine, et al., 1974) and that subjects' errors are systematic. That is, we assume that an analysis of cloze test errors (unacceptable responses) will reveal the strategies used by subjects in attempting to comprehend mutilated prose. (This latter assumption is, of course, supported by a long tradition of error analysis in L2 research: Corder, 1967; Richards, 1973; Dulay and Burt, 1972).

We found, as we began to analyze the data, that troublesome responses fell into two categories: responses which were clearly unacceptable but which indicated varying levels of comprehension, and responses which seemed acceptable but which were "unusual" or "non-native". In the first category are responses such as the following: (Most examples cited in this paper were produced by subjects participating in the study described in Clarke 1977. Words in parentheses are exact responses.)

1. Sometimes our __(sisters)__ ride in it too.
   a. sister
   b. car
2. It seldom rains and it __(never)__ snows.
   a. doesn't
   b. always
3. It (the food) was so good that I __(ate)__ too much during my visit.
   a. eat
   b. fat
4. And all this (the hotel accomodations) was __(not)__ expensive.
   a. very
   b. no
   c. small
   d. few

If the tests are strictly scored, none of the responses above can be given credit, and the test becomes more a measure of language proficiency than reading comprehension. Yet, the responses do indicate differences in comprehension. Responses 1b, 2b, and 4a are clearly wrong; the subject did not understand what information was required. With the other responses, however, it is not as easy to make a judgment. Responses 1a and 3a require only minor grammatical operations to become fully acceptable; sentence 2a requires only a minor change in the verb (3rd person singular to base) and none of the three (1a, 2a, 3a) would seriously affect communication if the sentences were produced in appropriate contexts. Response 4b shows the subject's confusion concerning the use of function words in the negative transformation, but the sentence would undoubtedly be correctly interpreted in conversation. Responses 3b and 4c,d seem to have some "lexical unity" with other words in the sentences in which they occur; if one eats too much one becomes fat, and "small" and "few" are appropriate if unacceptable expressions of degree, given the context.

In the second category (responses which are "non-native" or "unusual") we have:

5. I __(spent)__ a week in a hotel near the sea.
   a. stayed
   b. was
   c. had
   d. passed
   e. lived
   f. sheltered

Responses 5a-f all indicate comprehension. All are verbs, properly inflected, and all fall in the "semantic range" of words acceptable for describing a sojourn in a hotel. But the responses represent a continuum of acceptability, ranging from "perfectly acceptable" (stayed), through "acceptable, but causing a change in connotation" (was, had, passed) to "acceptable only in particular contexts" (lived, sheltered). The responses indicate receptive control of the language by the subjects; the differences among the responses reflect differences in the subjects' productive command of English.

In order to evaluate responses such as those in 1-5, we adapted an instrument originally developed by K. S. Goodman for the analysis of miscues in oral reading performance (Goodman and Burke, 1973, Appendix D). Each response is evaluated on three sets of criteria: syntactic acceptability (SYNAC), semantic acceptability (SEMAC), and semantic change (SEMCH). SYNAC and SEMAC provide an evaluation of the degree of acceptability of a response thus allowing us to rank unacceptable responses such as 1a-4d. SEMCH provides an estimation of the amount of meaning change caused by acceptable responses such as 5a-f. Codes and examples are given below:

Syntactic Acceptability (SYNAC)

4: totally acceptable:
  I ___stayed___ a week in a hotel by the sea.
  (spent)
3: acceptable in the sentence; the response satisfies sentence level syntactic constraints, but violates discourse constraints:
  After eating lunch I ___usually___ sleep for an hour.
  (would)
  (The passage requires the past tense.)
2: acceptable only with the following portion of the sentence; from the response on, the sentence is syntactically acceptable:
  The hotel food ___were___ very good.
  (was)
1: acceptable only with the preceding portion of the sentence; the sentence is syntactically acceptable up to and including the response:
  Sometimes our ___sister___ ride in it, too.
  (sisters)
0: totally unacceptable:
  It (the food) was so good that I ___fat___ too much during my visit.
  (ate)

Semantic Acceptability (SEMAC)

6: totally acceptable:
  I just wrote a ___hotel___ asking for a room in August.
  (letter)
5: totally acceptable if syntactic constraints are ignored; the sentence and/or response requires minor syntactic changes:
  Sometimes our ___sister___ ride in it, too.
  (sisters)
  It seldom rains and it ___doesn't___ snows.
  (never)
4: acceptable in the sentence; the response violates some passage-level meaning constraints:
  And all this (the hotel accomodations) was ___very___ expensive.
  (not)

3: acceptable in the sentence if syntactic constraints are ignored; the sentence and/or the response requires minor syntactic changes to become acceptable at the sentence level:
Even on Saturdays and Sundays she __don't__ work.
(must)

2: acceptable only with the following portion of the sentence: from the response on, the sentence is semantically acceptable:
At 12:00 I usually ate __speak__ with friends.
(lunch)

1: acceptable only with the preceding portion of the sentence; the sentence is semantically acceptable up to and including the response:
After eating __there__ big breakfast, I spent the morning swimming.
(a)

0: totally unacceptable:
The weather is wonderful __blue__.
(there)

9: indeterminate: the response seems to fit the context, but it is impossible to determine the contextual motivation for it:
It (the food) was so good that I __fat__ too much during my visit.
(ate)

If a response is coded SEMAC 6, indicating that it is totally semantically acceptable, it is further evaluated for semantic change.

Semantic Change (SEMCH)

3: no change; the exact response and the observed response are synonymous, given the context:
I __stayed__ a week in a hotel near the sea.
(spent)

2: minor change; there is a change in connotation:
I __was, had passed__ a week in a hotel near the sea.
(spent)

or, the response is logical, but changes the meaning of the sentence slightly:
I just wrote a __hotel__ asking for a room in August.
(letter)

or, no response is provided:
The hotel had a beautiful garden and a large __ø__ pool.
(swimming)

1: major change; the response conveys the author's meaning, but is "unusual" or "non-native":
I __sheltered__ a week in a hotel near the sea.
(spent)

Coding Conventions

In order to insure uniformity in coding, a number of conventions have been adopted.

Spelling errors. Spelling errors require coder judgment about the subject's intent. While this often involves "second guessing" the subject, coders can use evidence provided in the rest of the test to make their decisions.

6. Mrs. Brown was very __ungry__, but she thanked the policemen and they
(surprised)
left.

Whether *ungry* is coded as *angry* or *hungry* will depend on the subject's apparent language proficiency and reading ability as suggested by his responses to the other blanks. We make the assumption that a good reader is less likely to have produced *hungry* for this blank than a poor reader. Spelling errors which result in English words are coded as if the subject intended to write that word:

7. The hotel food ____tested____ very good.
(tasted)
8. How ____cold____ you get lost there?
(could)

This convention produces a conservative assessment of the subject's proficiency, thus reducing the amount of second guessing required of the coder. *Tested* receives codings of SYNAC 4, SEMAC 4; *cold* receives SYNAC 1, SEMAC 1. It is possible, of course, to rewrite the coding conventions so that phonetic spelling is taken into consideration. If, for example, Spanish is the native language of the subject who produced *tested*, full SEMAC credit could be justified on the grounds that he used Spanish phonology to spell an English word. A number of errors are probably not spelling errors at all, but rather results of learner strategies (Corder 1967, Richards 1973).

9. The ____childrens____ come home from school to eat lunch.
(children)
10. I ____thinked____ that he was a gentleman.
(thought)

The subjects who produced responses 9 and 10 understood what was required, and used the appropriate inflections. *Childrens* and *thinked* are defined as manifestations of interlanguage phenomena (Selinker 1972), and both are coded SYNAC 4, SEMAC 6.

Anomalous responses. We make the assumption that all responses (including non-responses and multiple word responses) are motivated by the subject's use of textual cues. Therefore, multiple words and blanks (ø) are evaluated using the same procedure as are one-word responses. With non-responses, it is necessary to include the word immediately preceding or following the blank when judging SYNAC and SEMAC.

11. The ____ø____ gentleman lost his way in the park.
(old)
12. By 8:00 the streets are again ____ø____ with people.
(crowded)

Sentence 11 is acceptable, of course, but 12 is coded SYNAC 4, SEMAC 2. In coding SEMAC, the sentence is read from *again;* ". . . again with people." is a possible structure ("The streets are crowded again with people."), so 12 is judged "acceptable with the following portion of the sentence". *Initial and final blanks:* When blanks appear at the end or beginning of a sentence, a response cannot receive SYNAC or SEMAC codes of 1 or 2. For obvious reasons, the following sentences must either be totally acceptable or totally unacceptable:

13. I guess that he was a gentleman although only a ____ø____.
(mouse)
14. ____ø____ she washes the dishes and cleans the kitchen.
(First)

Coding Instructions

It is, of course, impossible to completely separate the syntactic and semantic considerations of a piece of prose. But as sentences 1-4 show, responses can provide evidence of comprehension, yet be syntactically malformed. In order to assess the syntactic acceptability of a response independently of semantic acceptability, it is necessary to view each sentence in the abstract, considering the elements of the sentence and the obligatory features of those elements in the clause structure. To that end, each response is evaluated as to whether or not it fits syntactically into the clause structure of the sentence in which it occurs. The clause patterns of English, in their simple declarative form are:

A. A 2-element pattern
  1. SV The child was laughing.

B. Three 3-element patterns
  2. SVC Mary is kind.
  3. SVA Mary is here.
  4. SVO Somebody caught the ball.

C. Three 4-element patterns
  5. SVOC We have proved him wrong.
  6. SVOA I put the plate on the table.
  7. SVOO She gives me expensive presents.

(Quirk et al., 1972, p. 343)

Using these seven patterns as the acceptable clause structures occurring in English, the researcher can evaluate the syntactic acceptability of a response without reference to meaning. Each word is judged as a member of a constituent class (i.e.: transitive verb, plural noun, function word). The question to be asked by the coder is: "Are there *syntactic* features which forbid the occurrence of particular constituents in a particular clause pattern?" This procedure occasionally results in nonsense sentences receiving high syntactic coding:

15. We ____(work)____ at a gas station after school.
  a. go
  b. need

While neither 15a or 15b produces a meaningful sentence, only 15b is coded as syntactically malformed. The clause pattern is SVA(A); *go* fits the pattern $(\overset{S}{\text{We}}\ \overset{V}{\text{go}}\ \overbrace{\text{to the gas station}}^{A}\ \overbrace{\text{every day}}^{A}.)$, while *need* (being transitive) does not. The meaning (or lack of meaning) of a sentence is not taken into consideration; *go* is coded SYNAC 4 (totally acceptable) and *need* is coded SYNAC 1 (acceptable only with the preceding portion of the sentence).

Specifically excluded from SYNAC considerations are selection restrictions, that is, co-occurrence limitations imposed by the meaning of the word being evaluated:

Rules governing the kind of subject, object, or prepositional complement occuring

> with a particular verb come under the heading of selection restrictions. They differ from rules of concord, in that they do not involve two elements *sharing* the same feature, but one feature *projecting on to* another a feature which is necessary for its meaningful use.
>
> (Quirk et al., 1972, p. 371)

Sentences such as the following would receive SYNAC 4:

16. The boy dispersed.
17. The walls listened carefully.

Disperse requires a plural subject, but plurality here is a semantic requirement, not a syntactic one, as can be seen by the substitution of *crowd* (syntactically a collective singular) for *boy*. In sentence 17, the semantic feature of *+animate* is dictated by *listen*. In giving subjects syntactic credit for sentences such as 16 and 17, the assumption is made that they have recognized the structural requirements for a particular blank, but they have not yet mastered the rules governing co-occurrence restrictions.

In evaluating SEMAC, judgments of acceptability become more subjective, but partial objectivity is achieved by assessing meaning within a syntactic framework. While it could be argued that word order and expressive gestures are all that is required for communication in English, it is a fact that syntactic features carry a heavy semantic load in communication.[2] In addition, if one attempts to judge meaning without reference to syntax, the danger arises that the coding decisions will become so subjective as to preclude replication. For these reasons, the SEMAC coding for a particular response is not allowed to exceed the SYNAC coding. Two codes—SEMAC 3 and 5—designate responses which, except for minor syntactic error, would be coded 4 (acceptable in the sentence) or 6 (totally acceptable), respectively. Coder agreement on the definition of "minor" is crucial if any degree of objectivity is to be achieved. If coders are allowed too great a latitude in judging 3's and 5's, the situation is soon reached where more information is being provided by the researcher than by the subject. The definition of "minor syntactic error" will have to be determined within the objectives of each research project. The definition might include errors in person, tense, number, and use of function words:

18. We rides to school in it every day.
(ride)
19. Sometimes I play tennis.
(played)
20. The days are warm and sunny and the night are cool and clear.
(nights)
21. And all this (the hotel accomodations) was no expensive.
(not)

---

[2] Indeed, some researchers argue that meaning cannot be judged without well-formed syntax: "The structural organization of a sentence forms the basis for semantic relationships. Meaning . . . is dependent upon syntax. It is the order of the items and the use of inflection that indicate the meaning relationships of the items. The syntactic order is separate from and can precede the meaning but the meaning cannot exist without the order." (Goodman and Burke, 1973, Appendix D, p. 37) Most L2 teachers would probably feel that this statement is somewhat extreme given the communicative resourcefulness of many students.

There are few concrete instructions which can be suggested for improving the objectivity of the evaluation of SEMCH. Judgments of synonymity or "non-nativeness" are almost entirely idiosyncratic, and heavily influenced by one's dialect. Coders need to establish general guidelines, a task which is most easily accomplished through discussion of coding decisions. The system is not ideal, but it does allow researchers to arrive at an estimation of response appropriateness.

We have made an attempt to describe in detail the coding conventions and instructions necessary for effective use of this analysis tool. Of course, it is impossible to anticipate all of the coding problems that can occur in cloze response analysis, and there are several problems which will have to be solved within the framework of a given research project. One persistent problem is where to draw the line between "unacceptable" and "acceptable but non-native or unnatural". In our research, we have adopted a liberal definition of "acceptable" because we are interested in assessing receptive control of the language; for this reason the examples for SEMCH (page 136) contain responses which others might have coded as unacceptable.

Another problem concerns the analysis of responses involving function words. Often it is impossible to determine the amount of meaning change or loss caused by the deletion or insertion of a particle, determiner, preposition, etc. It is quite likely that the best solution is to adopt an arbitrary coding procedure and hope that the gains in scoring efficiency offset any loss in descriptive adequacy.

A final coding problem worth mentioning concerns the number of "minor syntactic errors" which can be ignored in awarding SEMAC codes 3 and 5. The solution to this problem also depends on the objectives of the research.

22. After eating lunch I ____(would)____ sleep for an hour.
a. went to
b. went
c. go to
d. go

Responses 22a-d indicate that the subject understood what the individual in the story did after lunch. *Went to* is fully acceptable, both structurally and semantically. The other three responses, however, contain minor syntactic errors (a deleted particle; an error in tense; both deleted particle and error in tense). The researcher must decide if equal credit for comprehension is to be given for 22b,c,d. In our judgment all three responses should receive codings of SEMAC 5, indicating a basic level of comprehension which is not significantly altered by syntactic error.

Discussion and Implications

The instrument presented here provides a framework within which cloze test responses can be objectively evaluated. Although a number of decision points require the judgment of a native (or near-native) speaker, it is hoped that the coding instructions are specific enough that a reference text (such as Quirk et al., 1972) can solve most coding problems. This scoring procedure requires a substantial investment of time and effort, but it provides a number of advantages over other scoring methods.

Most important is the fact that an objective assessment of acceptability facilitates replication of research. If a research report contains explicit definitions of acceptability, other researchers are more confident of their efforts at replication.

Furthermore, such a system permits researchers to determine their own criteria for acceptability. If one is interested in reading comprehension, for example, responses coded SEMAC 5 or better could be counted as acceptable.

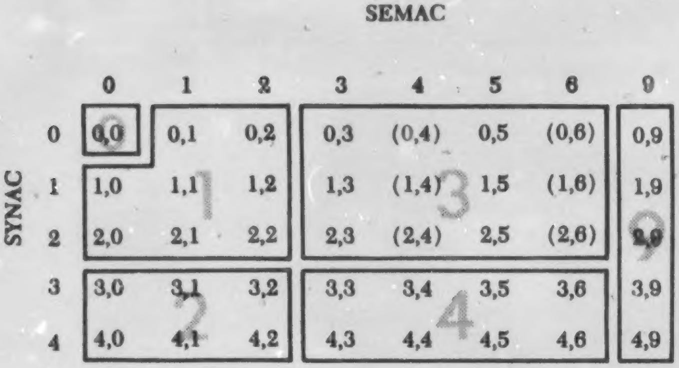
In addition, because SYNAC, SEMAC, SEMCH codings produce a profile of response quality for each subject, it becomes possible to conduct error analyses of cloze test performances. For example, an analysis could be conducted to determine the use of context by L2 learners. One might want to test the hypothesis that Group X (good readers, or high proficiency ESL students, or students from a particular language background, for example) makes more effective use of contextual clues than does Group Y (poor readers, or low proficiency ESL students, or ESL students from a different language background). Effective use of context is revealed in the subjects' SYNAC, SEMAC scores. Subjects who make effective use of context would be expected to receive a higher proportion of SYNAC 4's and 3's than SYNAC 2's, 1's, or 0's, and a higher proportion of SEMAC 6's, 5's, 4's, 3's than SEMAC 2's, 1's, 0's, or 9's.

The system also has value in L2 teaching, primarily as a reading diagnostic tool. A student's performance in cloze tests can be converted to a profile which indicates the percentage of responses in each code category. The information can be used by the teacher in preparation of materials or activities designed to solve problems students are having. A student whose responses are grammatically acceptable but nonsensical is concentrating too much on the task as a language test, and needs to be encouraged to attend to meaning, rather than form. A student whose answers show good comprehension but poor control of syntax could be provided with supplementary grammar work.

For teaching purposes, another code category can be developed by collapsing the data from SYNAC and SEMAC: context (CONTX). Context used by the subject to respond to the blank—coding for this category results from a consideration of SYNAC and SEMAC. (See Fig. 1.)

0: no discernible use of context
1: weak use of context
2: syntactic over semantic
3: semantic over syntactic
4: strong
9: indeterminate (automatic when SEMAC is 9)

CONTX permits the teacher to generalize about individual's performance on



SYNAC \ SEMAC	0	1	2	3	4	5	6	9
0	0,0	0,1	0,2	0,3	(0,4)	0,5	(0,6)	0,9
1	1,0	1,1	1,2	1,3	(1,4)	1,5	(1,6)	1,9
2	2,0	2,1	2,2	2,3	(2,4)	2,5	(2,6)	2,9
3	3,0	3,1	3,2	3,3	3,4	3,5	3,6	3,9
4	4,0	4,1	4,2	4,3	4,4	4,5	4,6	4,9

Figure 1. Coding for CONTX. (Because of the convention that semantic acceptability for a particular response cannot exceed syntactic acceptability, codes in parentheses cannot occur.)

cloze tests. Students whose responses are predominantly 0, 1, or 9 would benefit from basic language skills work. Students whose scores are predominantly CONTX 2 are too concerned with grammatical appropriateness, and should be encouraged to take risks, to guess, to concentrate on getting the meaning from a selection. From the point of view of reading instruction, students scoring 3's and 4's need only encouragement; their cloze test performance indicates that they are reading with understanding.

The analysis procedure described in this paper grew out of a need for an instrument which specified the linguistic criteria of "acceptability" in cloze test responses, and which revealed the processes used by subjects in responding to mutilated texts. By specifying the degree of syntactic and semantic acceptability and semantic change of each response the tool allows the researcher to develop profiles of subjects which can be used in testing hypotheses about ESL reading behavior, and in diagnosing learner difficulties in the ESL classroom. The description of code categories and the coding instructions were presented in detail with the hopes that other researchers would be able to use the tool in their own investigations.

REFERENCES

Berkoff, N. A. 1976. A diagnostic use of cloze testing. *English Teacher's Journal* (Israel) *6*, 42–49.

Bowen, J. D. 1969. A tentative measure of the relative control of English and Amharic by 11th grade Ethiopian students. *UCLA TESL Workpapers 2*, 69–89.

Clarke, M. A. 1977 A psycholinguistic description of the reading performance of selected adult native Spanish speakers reading in Spanish and English: An investigation into reading universals. Ph.D. Dissertation, University of Michigan. (in preparation).

Corder, S. P. 1967. The significance of learners' errors. *IRAL 5*, 4, 161–169.

Darnell, D. K. 1970. Clozentropy: a procedure for testing English language proficiency of foreign students. *Speech Monographs 37*, 1, 36–46.

Dulay, H. C., and M. Burt. 1972. Goofing: an indication of children's second language learning strategies. *Language Learning* 22, 2, 235–252.

Eskey, D. E. 1973. A model program for teaching advanced reading to students of EFL. *Language Learning* 23, 2, 169–184.

Goodman, K. S. 1970. Reading: A psycholinguistic guessing game. In H. Singer, and R. B. Ruddell (eds.) *Theoretical models and processes of reading*, Newark, Del, International Reading Association. Pp. 259–271.

Goodman, K. S., and C. Burke. 1973. *Theoretically based studies of patterns of miscues in oral reading performance*. U.S. Department of Health, Education and Welfare, Office of Education.

Irvine, P., J. W. Oller, P. Atai. 1974. Cloze, dictation, and the Test of English as a Foreign Language. *Language Learning 24*, 2, 245–252.

Oller, J. W. 1972. Scoring methods and difficulty levels for cloze tests of ESL proficiency. *Modern Language Journal 56*, 3, 151–158.

Oller, J. W., and C. A. Conrad. 1971. The cloze technique and ESL proficiency. *Language Learning 21*, 1, 183–196.

Oller, J. W., D. Bowen, T. T. Dien, and V. W. Mason. 1972. Cloze tests in English, Thai, and Vietnamese: native and non-native performance. *Language Learning 22*, 1, 1–16.

Plaister, T. 1973. Teaching reading comprehension to the advanced ESL student using the cloze procedure. *RELC Journal 4*, 31–38.

Quirk, R., S. Greenbaum, S. Leech, and J. Svartvik. 1972. *A grammar of contemporary English*. New York, Seminar Press (Harcourt Brace, Jovanovich, Inc.).

Rankin, E. J. 1957. An evaluation of the cloze procedure as a technique for measuring reading comprehension. Ph.D. Dissertation, University of Michigan.

Richards, J. C. 1973. Error analysis and second language strategies. In J. W. Oller, and J. C. Richards (eds.) *Focus on the learner: pragmatic perspectives for the language teacher*, Rowley, Mass., Newbury House.

Selinker, L. 1972. Interlanguage. *IRAL 10*, 3, 209–231.

Stubbs, J. B., and G. R. Tucker. 1974. The cloze test as a measure of ESL proficiency for Arab students. *Modern Language Journal* 58, 239–241.

Taylor, W. L. 1953. Cloze procedure: A new tool for measuring readability. *Journalism Quarterly 30*, 415–433.

Part 3

Learning: Research in Second Language Acquisition

Some Issues Relating to the Monitor Model

Stephen D. Krashen

University of Southern California

This paper is a discussion of several issues relating to the "Monitor Model" for second language performance. The Monitor Model claims that adults can both subconsciously *acquire* and consciously *learn* second languages, and that learning is available only as a Monitor, or as an editor of the output of the acquired system. The following issues are briefly discussed: 1) The possibility that acquisition may also be used as a monitor. 2) The conditions for Monitor use. It is proposed that performers monitor when they have sufficient processing time and when they are "focussed on form". 3) Certain results in studies of grammatical morphemes. It is claimed that the adult "natural order" for grammatical morphemes is found in un-monitored conditions. The natural order is not due to the use of one particular instrument or modality, and when a sufficiently large data base is used (a minimum of ten obligatory occasions per morpheme) there is surprisingly little individual variation and good agreement between longitudinal and cross-sectional studies. 4) Limits on learning. Learning is limited in three ways: It is not available to all adult performers, it is in general applicable to "easy" rules, and it is only available in certain situations. 5) Need learning precede acquisition? It is suggested that while such a progression may appear to occur, it is not necessary and is not an underlying process. 6) The role of the first language. The first language may serve as a substitute utterance initiator in situations where little acquisition has taken place (e.g. foreign language as opposed to second language situations).

This paper is devoted to a discussion of some issues that relate to the "Monitor Model", a general model for adult second language performance. I will not attempt a complete discussion of all aspects of the model, as details are available in several places in the literature (Krashen 1976, 1977) and more will be appearing soon (Krashen, in press). Instead, I will focus on some specific points which represent areas that may have been vague and where newer developments have clarified the interaction between acquisition and learning in the adult second language performer.

The principal claim of the Monitor Model is that the adult second language performer has two means for internalizing rules of the target language. One is language *acquisition*, which is very similar to the creative construction process utilized in both child L1 and child L2 acquisition. Research in child language acquisition has taught us that acquisition is primarily subconscious, not influenced by overt teaching, and is encouraged by simplified input and participation in conversation (for reviews, see Cazden, 1972; Brown, 1973; Clark and Clark, 1977). The other process available to many adults is language *learning*, which results in conscious representation of pedagogical rules. Overt teaching is thought to be quite useful for learning (Krashen and Seliger, 1975). The Monitor Model hypothesizes that learning is available to the adult performer in second language production only as a Monitor, that is, we can use conscious grammar only to alter the output of the acquired system.

The acquisition-learning distinction, as I have outlined it, is not new: Lawler and Selinker (1971) propose that for rule internalization one can "postulate two distinct types of cognitive structures: (1) those mechanisms that guide 'automatic' language performance . . . that is, performance . . . where speed and spontaneity are crucial and the learner has no time to consciously apply linguistic mechanisms . . . and (2) those mechanisms that guide puzzle- or problem-solving performance . . ." (p. 35). Corder (1967), citing an unpublished paper by Lambert, also discusses the acquisition-learning distinction and the possibility that acquisition is available to the adult second language performer. The Monitor Model makes some very specific hypotheses about the inter-relation between acquisition and learning in the adult, with the main hypothesis being that learning is used only as a Monitor. In other papers I have argued that this hypothesis sheds light on nearly every issue under discussion in second language theory and practice.

One point that has caused some confusion is the fact that the model allows for acquisition to serve as a Monitor as well. When native speakers self-correct in their first language, as in cases of speech errors, conscious rules are usually not called upon. Rather, the utterance is re-initiated via the acquired system. This may also happen with Monitor "under-users", those who generally do not use conscious rules in second language performance under all conditions (Krashen, in press). Again, the Monitor Model states that learning is only used as a Monitor. Learning does not initiate performance or underlie second language fluency in adults.

Conditions for Monitor Use

A second point concerns the conditions for Monitor use. In earlier papers, it was hypothesized that learning was available only when the performer had *time*, as in written performance: In situations where little processing time is available, the model predicts that the performer will rely primarily on the acquired system. Again, the Monitor Model is not the first to propose this constraint on the application of conscious rules, as indicated by the Lawler and Selinker quote given above (see also Lawler and Selinker, 1971, p. 38).

Marina Burt and Heidi Dulay have pointed out to me (personal communication; see also Dulay and Burt, 1977), however, that this condition for Monitor use is not sufficient: Slow performance is not always Monitored. Rather, it is hypothesized that Monitoring occurs when the performer is *focussed on form*, when he or she is concerned with the "correctness" of the utterance as well as (or instead of) its communicative function. As we shall see below, Burt and Dulay's insight clarifies operation of the Model considerably.

Studies of Grammatical Morphemes

Another point that deserves some discussion is the current state of the art in the study of the acquisition and difficulty orders for grammatical morphemes and the relation of these studies to the Monitor Model. In previous papers, it was suggested that the Monitor Model successfully predicted the variation in performance on grammatical morphemes seen in adult studies: When adult ESL performers produce English under "monitor-free" conditions (little processing time and focus on communication rather than on just form), their difficulty order for grammatical morphemes is similar to that seen in child second language performance (Dulay and Burt, 1975). The appearance of the child's difficulty order under these conditions is hypothesized to be the manifestation of the creative construction process in adults (discussion in Krashen, 1977, in press). When pencil and paper "grammar" type tests are used, adult performers can focus on form and have time to think about specific rules, and the "natural order" of grammatical morphemes is disturbed. Specifically, morphemes that are generally acquired late (not yet learned) and are "easy" to conceptualize, or learn, rise in rank.

Is the Natural Order tied to the BSM?

Some objections have been raised to this picture. First, it has been suggested that the adult natural order is somehow tied to the use of the Bilingual Syntax Measure (Murt, Dulay, and Hernandez, 1975) as an elicitation device, as the BSM was used in the Dulay and Burt child second language studies and in the Bailey, Madden, and Krashen (1974) and Larsen-Freeman (1975) adult studies. We recently tested this possibility by determining grammatical morpheme difficulty order using adult free (unmonitored) speech (Krashen, Houck, Giunchi, Bode, Birnbaum, and Strei, 1977). The results, given in Table 1, show clearly that the adult "natural order" is present here as well: Our difficulty order correlates quite well with the order obtained using the BSM for both children and adults. Also, as I shall describe below, we obtained similar results using composition, as has Andersen (1975).

Individual Variation: Cross-Sectional versus Longitudinal

It has also been suggested that there is considerable individual variation in morpheme orders, and that longitudinal and cross-sectional studies do not always agree (Rosansky, 1976). In an attempt to determine just how much variation really exists, I recently reviewed every study available to me where grammatical morphemes were analyzed in obligatory occasions. This included child L1, child L2, delayed L1, and adult agrammatics. It included both grouped and individual studies, and longitudinal and cross-sectional studies. The complete list of reports consulted is given in Table 2.

Following de Villiers (1974) I only included morphemes with at least ten obligatory occasions in a given study. This is an extremely small number, and I originally thought that this would produce large variation; just one additional correct response, for example, would change a subject's score on a morpheme

TABLE 1

Grammatical Morphemes in Adult Free Speech
(from Krashen, Houck, Giunchi, Bode, Birnbaum, and Strei, 1977)

Morpheme	Accuracy
copula	.87
ing	.84
s. plural	.71
article	.69
irr. past	.67
reg. past	.64
auxiliary	.56
III singular	.36

Correlations of free speech order with other studies

Study	Elicitation device	Correlation (rho)
Bailey et al., 1974	BSM	.857**
Larsen-Freeman, 1975	BSM	.690*
Dulay and Burt, 1974	BSM	.762*

* $p<.05$

** $p<.01$

TABLE 2

Studies Analyzing Grammatical Morphemes in Obligatory Occasions

I. Child L1 acquisition
- A. Individual cross-sectional: None
- B. Individual longitudinal: Brown, 1973 (Adam, Eve, Sarah), Curtiss, Fromkin, and Krashen, 1975 (Genie).
- C. Grouped cross-sectional: de Villiers and de Villiers, 1973. Kessler, 1975.
- D. Grouped longitudinal: Brown, 1973.

II. Child L2 acquisition
- A. Individual cross-sectional: Kessler and Idar, 1977 (Than, three cross-sections). Rosansky, 1976 (Jorge, 10 cross-sections, Marta, Cheo, Juan).
- B. Individual longitudinal: Rosansky, 1976 (Jorge), Hakuta, 1974 (Uquisu), Kessler and Idar, 1977 (Than).
- C. Grouped cross-sectional: Dulay and Burt, 1973; Dulay and Burt, 1974.

III. Adult L2 acquisition
- A. Individual cross-sectional: Birnbaum, 1976 (Hector), Strei, 1976 (Gerardo), Holdich, 1976 (Holdich), Rosansky, 1976 (Alberto, Dolores).
- B. Individual longitudinal: none.
- C. Grouped cross-sectional: Bailey, Madden, and Krashen, 1974. Larsen-Freeman, 1975. Krashen, Houck, Giunchi, Bode, Birnbaum, and Strei, 1977. Andersen, 1976. Birnbaum, Butler, and Krashen, 1977 (also in Krashen, Butler, Birnbaum, and Robertson, 1976).

IV. Agrammatics
- A. Individual cross-sectional: de Villiers, 1974 (A3,50,14,5,24,6,43 (2 cross-sections))
- B. Individual longitudinal: none.
- C. Grouped cross-sectional: de Villiers, 1974.

by ten percentage points. Nevertheless, I found an amazing amount of uniformity across all studies that used monitor-free instruments, that is, in all studies where language was used for communication. It should be noted that Rosansky (1976) allowed items to be analyzed that appeared in less than ten obligatory occasions for an individual cross-section. This probably accounts for our different conclusions, which are based on some of the same data.

Table 3 lists the particular relations discovered and describes the criteria

TABLE 3

Ordering Relations Among Grammatical Morphemes

I. Clear cases (True at least 90% of the time. Data from studies where each morpheme occurs in at least ten obligatory occasions. Counter-examples discussed in table four.)

1. ING precedes Irr. PAST
2. ING precedes Reg. PAST
3. ING precedes III singular
4. ING precedes POSS
5. ING precedes AUX
6. PLURAL precedes Irr. PAST
7. PLURAL precedes Reg. PAST
8. PLURAL precedes III singular
9. PLURAL precedes POSS
10. COP precedes Irr. PAST
11. COP precedes Reg. PAST
12. COP precedes III singular
13. COP precedes POSS
14. ART precedes POSS
15. ART precedes III singular
16. ART precedes Reg. PAST
17. AUX precedes POSS
18. AUX precedes III singular

(1 through 9 hold for all studies; 10 through 18 do not necessarily hold for child L1 acquisition)

II. Less clear cases (true of most studies where each morpheme occurs in at least ten obligatory occasions. Except where indicated, probability of relation holding significant by sign test).

1. ING precedes ART
2. ING precedes PLURAL
3. PLURAL precedes AUX
4. PLURAL precedes ART ($p<.15$)
5. COP precedes AUX
6. COP precedes ART
7. ART precedes Irr. PAST
8. AUX precedes Irr. PAST
9. AUX precedes Reg. PAST
10. Irr. PAST precedes III singular
11. Irr. PAST precedes Reg. PAST ($p<.21$)
12. Irr. PAST precedes POSS ($p<.11$)

(Numbers 1, 3, and 7 hold for child L1 acquisition; the others do not necessarily hold for child L1 acquisition.)

---

for selection. The counter-examples to these generalizations (Table 4) are rarely "serious": As table four indicates, they often fall within ten percentage points, are items represented by less than 20 obligatory occasions, or are within one or two ranks where percentages are not given, or come from Uguisu, Hakuta's subject. (Detailed discussion of Uguisu will wait for another paper.)

TABLE 4

Counterexamples to Ordering Relations among Grammatical Morphemes

total instances consistent with relations given in table three	= 881
total instances inconsistent with relations given in table three	= 98 (10%)
Analysis of counterexamples:	
a. within ten percentage points	= 49 (50%)
b. within twenty percentage points	= 9
c. where percentages not given: one rank difference	= 16
d. where percentages not given: two rank difference	= 9
e. studies producing "true" counterexamples (not a-d above):	
1) Uguisu (Hakuta, 1974)	= 5
2) Larsen-Freeman (Imitation I)	= 2
3) Jorge 11 (Rosansky, 1976)	= 1
4) Dolores (Rosansky, 1976)	= 2
5) Andersen, 1976[a]	= 1
6) agrammatic individual subjects (de Villiers)[b]	= 4

[a] number of obligatory occasions not known

[b] obligatory occasions known to be more than 10 but may be less than twenty

Many of the generalizations hold for L1 as well as L2. In general, this is true of bound morphemes, and this agreement underlies the consistently positive but often statistically insignificant correlation one sees between L1 and L2 scores when rank order correlations are used.

Figure 1 gives a proposed "natural order" for second language and for

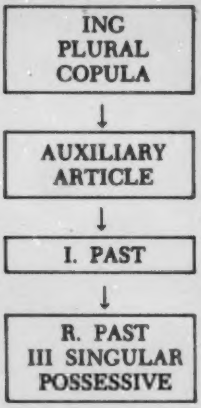


Figure 1. Proposed "Natural Order" for Second Language Acquisition and Agrammatics. (No claims are made about ordering relations for morphemes in the same box)

Broca's aphasia, and was constructed from the relations given in table three. It is extremely interesting to note that all "Monitor-free" studies using adult subjects (except Imitation I and II; see Larsen-Freeman (1977) for discussion of Imitation I), show a healthy positive correlation with this order.[1] (Table 5)

Looking at morphemes in obligatory occasions does not give us a complete picture of the processes underlying language acquisition (see Anderson, 1977,

TABLE 5

Rank Order Correlations with the Proposed "Natural Order"

I. Studies with nine countable grammatical morphemes:

study[a]	*rank order correlation*[b]
Marta	.823
Uguisu	.170
Dulay and Burt, 1974 – combined	.770
Dulay and Burt, 1974 – Spanish	.753
Dulay and Burt, 1974 – Chinese	.827
Dolores	.849
Andersen, 1975	.883
Larsen-Freeman, BSM time I	.840
Larsen-Freeman, BSM time II	.857
Larsen-Freeman, Imitation I	.416
Larsen-Freeman, Imitation II	.719
Larsen-Freeman, Listen I	.719
Larsen-Freeman, Listen II	.606
Larsen-Freeman, Read I	.130
Larsen-Freeman, Read II	−.130
Larsen-Freeman, Write I	.217
Larsen-Freeman, Write II	.148

for n = 9, significance at .05 level requires rho = .600 or larger.
significance at .01 level requires rho = .783 or larger.

---

[1] Kendall W (coefficient of concordance) was computed for those studies containing the same morphemes in the minimum number of obligatory occasions. For studies with nine morphemes in common (Marta, Uguisu, Dulay and Burt, 1974-combined, Dolores, Andersen, 1977, and Larsen-Freeman's two administrations of the BSM), W=.619, p<.001. For studies with the same eight morphemes in common (Jorge 7, 11, 18, 20, Krashen et al., 1977), W = .64, p<.01. For studies with the same seven morphemes in common (Birnbaum, Butler, and Krashen, 1977 – Free I, Edit I, Free II, Edit II, Cheo, Alberto). W = .618, p<.01. de Villiers (1974) computed a Kendall W of .60 (p<.001) for her individual agrammatic subjects.

The relationships proposed in figure one are also supported in Andersen (1977), who reanalyzed his 1975 data in several interesting ways. Andersen also presents data indicating significant agreement among individual subjects. Additional evidence against excessive individual variation is Bailey et al. (1974), who found "a high level of agreement" among different classes of ESL students for grammatical morpheme difficulty order. Each subgroup contained about ten students.

While all correlations with the "natural order" for Monitor-free studies are positive, a few miss statistical significance at the .05 level. This is occasionally due to unusual performance in one morpheme: In Juan, for example, there was very high performance in the III singular morpheme (16/16). In my judgment, this failure to reach significance in every case is not serious, as several studies that "miss" come quite close (e.g. Cheo) and the effect is reliable. See Ferguson (1971), among others, for a discussion of the prevalence of type II errors when such near misses are analyzed as non-rejection of the null hypothesis when n's are small, as they are here.

TABLE 5 *(Cont.)*

II. Studies with eight countable grammatical morphemes:

study	*rank order correlation*
Jorge 7	.878
Jorge 11	.615
Jorge 18	.725
Jorge 20	.886
Dulay and Burt, 1973 – Sacramento	.939
Dulay and Burt, 1973 – East Harlem	.878
Dulay and Burt, 1973 – San Ysidro	.865
Bailey, Madden, and Krashen, 1974	.939
Krashen et al., 1977	.779

for n = 8, significance at .05 level requires rho = .643 or larger.
significance at .01 level requires rho = .833 or larger.

III. Studies with seven countable grammatical morphemes:

study		*rank order correlation*
Cheo		.712
Juan		.368
Jorge 9		.769
Jorge 13		.862
Jorge 15		.726
Holdich		.726
Alberto		.730
Birnbaum et al., Free I		.726
Birnbaum et al., Edit I		.802
Birnbaum et al., Free II		.557
Birnbaum et al., Edit II		.712
de Villiers, 1974 – agrammatics combined		.880
agrammatics:	A3	.955
	A50	.599
	A14	.962
	A43 time I	.749
	A43 time II	.755
	A5	.637
	A24	.768
	A6	.805

for n = 7, significance at .05 level requires .714 or larger.
significance at .01 level requires .893 or larger.
(one-tail)

---

[a] All individual cross-sections listed, with the exception of Uguisu (Hakuta, 1974) and Holdich (Holdich, 1976), and the agrammatics, are from Rosansky (1976).

[b] All correlations have been corrected for ties.

Wagner-Gough and Hatch, 1976). Nevertheless, the data presented here strongly confirms the reality of a "natural order", a reliably occuring order in longitudinal and cross-sectional, individual and grouped studies of second language performers. Admittedly, it is not a rigidly invariant order, as shown in figure one. It is also far from random, however.

The task is far from complete in the area. As indicated in table two, more data should be gathered for adults longitudinally, and the effect of Monitoring on the morpheme order deserves replication.

How Important is Learning?

Another central issue concerns the relative importance of learning and acquisition. As stated, the model clearly implies that acquisition is central and learning, the Monitor, is peripheral. We now have several sorts of evidence that confirm this position, evidence that *learning* is *limited* in at least three important ways: with respect to *individuals* (not everyone uses the Monitor in performance, even when time is available and they are focussed on form), with respect to *rules* (learning may be applicable only to a small subset of the grammar), and with respect to *situations* (it is not used everywhere).

Monitor Under-users

Some recent case histories have indicated that many ESL performers are Monitor "under-users", that is, they do not utilize conscious rules under any conditions. Some of these performers simply do not know the rules, while others, for some reason, just don't use them. In Lawler and Selinker's terms, this includes "individuals who cannot learn a particular verbalized rule . . . (and) . . . individuals who can learn such a rule and not be able to apply it with or without time to do so" (p. 39). A particularly interesting phenomenon is the performer who pays lip-service to the value of conscious rules but who nevertheless does not use them. Stafford and Covitt (in press) have pointed this out in a recent paper, and present the case of "V", an ESL student who claimed that "grammar is the key to every language" on a questionnaire. When self-correcting, however, V admitted that he appeals to his "feel" for English and does not use conscious rules because he simply doesn't know any. (For more case histories, see Krashen, in press, and Stafford and Covitt, in press).

Limitations on Rule Learning

A second limitation on learning is with respect to rules. Our current hypothesis is that the Monitor is limited to items that (1) have not yet been acquired, and are (2) in a sense "easy" to carry around as mental baggage. Our composition study, reviewed below, is consistent with this, as are case histories and error analyses. Errors on such late acquired (in the technical sense) but easily conceptualized items as the third person singular ending and regular past ending are typically described as "careless". Students report that they consciously "know" the rule but carelessly fail to apply it (Dušková, 1969; Cohen and Robbins, 1975; Krashen and Pon, 1975). Our subject "P", for example, (Krashen and Pon, 1975), a linguistics major at Queens College, wrote nearly error-free English, but omitted such "easy" late morphemes in casual speech.

Difficult rules, it is proposed, are *acquired* if they are to be performed correctly in all situations. Difficulty may be defined in terms of the semantics or pragmatic conditions for rule application (as in certain aspects of article use;

see e.g. Linde and Labov., 1975, or Lackstrom, Selinker, and Trimble, 1973) and/or in terms of the syntactic operations required, the mental gymnastics, for rule application. This latter kind of difficulty is illustrated by the inability of even fairly advanced adult second language performers in French to produce sentences like

(1) Je ne lui en ai pas donné.
I didn't give any of it to him.

with any fluency, as such sentences require a great deal of movement and ordering of elements. According to texts and syllabi, such sentences should be within the capacity of second semester students of French as a foreign language.

Another confirmation of the hypothesis that adults may only in general apply conscious learning to certain easy rules in performance is Schlue's recent MA Thesis (Schlue, 1977). Her subjects were three intermediate level ESL students at UCLA. Here, "intermediate" means a foreign student who is studying ESL but who is also taking a partial academic load. Such students have "been through" the grammar, having completed texts such as Praninskas' *Rapid Review of English Grammar*. In other words, they are expected to "know" the grammar. Schlue tape recorded conversations with these students and asked them to correct their own errors from the tape. They were thus encouraged to focus on form and supposedly tapped the Monitor. The students caught about 30% of their errors this way, or, stated differently, over ⅔ of their errors got through! Thus, most errors survived, even when the students were encouraged to focus on form and apply conscious rules. An interpretation of this result is that the students simply did not know enough rules of the target language; their monitored, or learned, competence was only a small portion of what they needed to correct most errors.

When is Learning Available?

The third limitation on learning is *situation*. Our current data confirms that the conscious grammar is most available on grammar tests, is hardly available at all in oral communication, and is only used to a modest degree in composition. This is predictable from the two conditions on Monitor use given above. Subjects have the most time in grammar tests and are deliberately focussed on form. Oral speech is usually the mirror image of this situation; there is little processing time and communication is usually central. The composition situation illustrates that the focus on communication condition is more potent than the lack of time condition for Monitor-free performance. Here, performers often have time, but learning makes a modest contribution to performance in writing. Our recently completed study (Birnbaum, Butler, and Krashen, 1977; see also Krashen, Butler, Birnbaum, and Robertson, 1976) confirms this. We asked ESL students to write both free (write as much as you can in ten minutes) and edited (concentrate on correct spelling and grammar) compositions. We found

(1) the morpheme "natural order" is seen in both free and edited writing, an order quite close to that found in oral studies, and

(2) the difference in accuracy between the free and edited tasks, for the six morphemes examined, was quite small (about six percent). The only consistent change was a clear but unspectacular increase in the third person singular morpheme, which was not enough to disturb the natural morpheme order. Accuracy was thus only modestly affected by editing.

Schlue's data also speaks to this point. She reported that self-corrections in free speech (non-induced corrections) for her three foreign students, while helpful in increasing accuracy, only occured on 7.2% of all errors.

From this kind of data we can conclude that when real communication, or natural language use is involved, learning is not fully available. Note also that the Schule and composition studies probably overestimate Monitor use, as some of the self-correction may have been done on the basis of acquisition, or by a "feel" for correctness. Combined with the two other Monitor limitations discussed above, this result implies that the Monitor is used only by some people, for some rules, and for more artificial language use situations. It is interesting that some approaches to second language teaching presume nearly the opposite, that conscious rules are utilized by all performers, for all rules, and in all performance.

This is not to say that learning is evil and should not be appealed to at all. Very successful second language performers often have an extensive conscious knowledge of the target language in addition to advanced acquisition. They appeal to the Monitor, however, when Monitor use does not interfere with communication, and use the Monitor as a supplement to correct the output of their acquired competence. They are thus able, like "P" in Krashen and Pon (1975) to achieve the illusion of native speaker competence in certain situations, as in some kinds of writing and prepared speech.

Does Learning have to precede Acquisition?

A corollary of the above discussion of the centrality of acquisition is that acquisition need not be preceded by learning. While such a progression often appears to exist, several facts speak against a conversion of learning to acquisition as an underlying process in the development of fluency. First, we see an impressive range of syntactic mastery in many non-Monitor users (see the case history of Hung in Cohen and Robbins (1976) and other case histories in Stafford and Covitt, in press). Second, the limitation of learning to a limited number of grammatical rules also implies that adults must acquire some items right from the start. Third, teachers sometimes report that after second language students are "taught" a complex rule, they discover that they had a "feel" for the structure already. Students are often pleased to obtain confirmation for this tacit knowledge and feel that they have learned something important, but it is questionable whether they have increased their capacity to perform in

English by such conscious rule learning. Finally, if learning had to precede acquisition, how could we explain the failure of even advanced students to use certain items correctly under all conditions? "P", for example, knew the third person singular rule as well as anyone and had had years of practice with it. It had been learned, but not acquired, a good case of a rule that is late-acquired (low in both difficulty and acquisition orders in the studies listed in table two) and easy to conceptualize and learn. All this data and these observations become clear when acquisition and learning are considered to be independent, each with it's own processes.

The Role of the First Language

Previously published work on the Monitor Model has also neglected the relationship of the first language to acquisition and learning (for brief remarks, see Krashen, in press). The first language fits neatly into the Monitor Model, and it's role becomes clear when earlier results and observations on first language influence are considered. I summarize the relevant findings here:

(1) First language influence appears to be strongest in complex word order and in word for word translations of phrases (LoCoco, 1975; Politzer, 1968; Duškova, 1969; Richards, 1975; Grauberg, 1968).

(2) L1 influence is weaker in bound morphology (references as in (1) above).

(3) L1 influence appears to be strongest in *acquisition-poor* environments. Dulay and Burt (1974) have demonstrated that L1 influence is rare in child second language acquisition (but see below). On the other hand, studies that report a high amount of first language influence, such as those cited in points (1) and (2) above, are mostly *foreign* and not second language studies, situations in which natural appropriate intake is scarce and where translation exercises are frequent. In this regard, it is interesting to note that we can find signs of first language influence in immersion bilingual programs where input is often primarily from the teacher and not from peers. First language influenced errors here are also in the domain of word order (Plann and Ramirez, 1976; Selinker, Swain and Dumas, 1975).

(4) Consider Newmark's proposal for a mechanism for L1 influence (Newmark, 1966): L1 influence is not proactive inhibition, but is simply the result of the performer being "called on to perform before he has learned the new behavior". The result is "padding", using old knowledge, supplying what is known to make up for what is not known". Newmark suggests that the "cure for interference is simply the cure for ignorance: learning" (in terms of the Monitor Model, this would read "acquisition").

What can be concluded from the above is that the L1 may "substitute" for the acquired L2 as an utterance initiator when the performer has to produce in the target language but has not acquired enough of the L2 to do this.

It may in fact be the case that the hypothesized domain of acquisition in L2 performance is the same as those rules that are most prone to L1 influence, while aspects of the target language that may be learned (late acquired, easy to conceptualize; e.g. bound morphology) are relatively free of L1 influence.

First language influence may therefore be an indication of low acquisition. If so, it can be eliminated or at least reduced by natural intake and language use. This is what apparently occured in Taylor's ESL subjects, who showed less first language influence with more proficiency (Taylor, 1975).

Perhaps the "silent period" observed in natural child second language acquisition (Hakuta, 1974; Huang and Hatch, in press) corresponds to the use of the L1 in un-natural adult second language performance. The children may be building up acquired competence via input, and several recent studies (Gary, 1974; Postovsky, 1977) imply that less insistence on early oral performance may be profitable for children and adults.

First language influence can thus be considered as un-natural. One could theoretically produce sentences in a second language without any acquisition: The first language surface structure can be used with second language content lexicon inserted. The Monitor may then be used to add some morphology and do its best to repair word order where it differs from the L1. One can only go so far with this mode, as one is limited by the competence of the conscious grammar and one must appeal to it with every utterance. The adult can, however, produce sentences right away in the target language using this mode, and this may account for reports of more rapid progress in early stages for adults than for children in second language performance (Snow and Hoefnagel-Hohle in press). It is a temporary advantage, however. Acquisition may be slow, but it is, in the long run, much more useful when language is used for the purpose of communication. After all, that's what language is for.

REFERENCES

Anderson, R. 1976. A function acquisition hierarchy study in Puerto Rico. Paper presented at the 10th Annual TESOL Conference, New York. March, 1976.

Anderson, R. 1977. The impoverished state of cross-sectional morpheme acquisition/accuracy methodology. Paper presented at the Los Angeles Second Language Research Forum, UCLA. February, 1977.

Bailey, N., C. Madden, and S. Krashen. 1974. Is there a "natural sequence" in adult second language learning? *Language Learning 24*, 235-243.

Birnbaum, R. 1976. Transcription and analysis of the speech of an adult second language learner. Term paper, Department of Linguistics, USC.

Birnbaum, R., J. Butler, and S. Krashen. 1977. The use of the Monitor in free and edited ESL compositions. Paper presented at the Los Angeles Second Language Research Forum, UCLA. February, 1977.

Brown, R. 1973. *A first language*. Cambridge, MA., Harvard Press.

Burt, M., H. Dulay, and Ch. E. Hernandez. 1975. *The bilingual syntax measure*. New York, Harcourt Brace Jovanovich.

Cazden, C. 1972. *Child language and education*. New York, Holt, Rinehart, and Winston.

Clark, H., and E. Clark. 1977. *Psychology and Language*. New York, Harcourt Brace Jovanovich.

Cohen, A., and M. Robbins. 1976. Toward assessing interlanguage performance: The relationship between selected errors, learner's characteristics, and learner's explanations. *Language Learning 26*, 45-66.

Corder, S. P. 1967. The significance of learner's errors. *IRAL 5*, 161–170.

Curtiss, S., V. Fromkin, and S. Krashen. 1975. Language development in the mature (minor) right hemisphere. Paper presented at the Linguistics Society of America, December, 1975. Also in press, *ITL: Review of Applied Linguistics*.

de Villiers, J., and P. de Villiers. 1973. A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research 2*, 267–278.

de Villiers, J. 1974. Quantitative aspects of agrammatism in aphasics. *Cortex 10*, 36–54.

Dulay, H., and M. Burt. 1974. Errors and strategies in child second language acquisition. *TESOL Quarterly 8*, 129–136.

Dulay, H., and M. Burt. 1975. A new approach to discovering universal strategies of child second language acquisition. In D. Dato, (ed.) *Developmental psycholinguistics: theory and applications*, Georgetown University Roundtable on Languages and Linguistics 1975. Washington, D.C., Georgetown University Press. Pp. 209–233.

Dulay, H., M. Burt. 1977. Remarks on creativity in language acquisition. In M. Burt, H. Dulay, and M. Finocchiaro (eds.) *Viewpoints on English as a second language*, New York, Regents. Pp. 95-126.

Dušková, L. 1969. On sources of error in foreign language learning. *IRAL 4*, 11–36.

Ferguson, G. 1971. *Statistical Analysis in Psychology and Education*. New York, McGraw-Hill.

Gary, J. O. 1974. The effects on children of delayed oral practice in initial stages of second language learning. Ph.D. dissertation, UCLA.

Grauberg, W. 1971. An error analysis in German of first-year university students. In G. Perren and J. Trim (eds.), *Applications of Linguistics*, Cambridge, England: University Press. Pp. 257-263.

Hakuta, K. 1974. A preliminary report on the development of grammatical morphemes in a Japanese girl learning English as a second language. *Working Papers on Bilingualism* 3, 18–43.

Huang, J., and E. Hatch. A Chinese child's acquisition of English. In E. Hatch (ed.), *Second Language Acquisition*, Rowley, Newbury House. In press.

Holdich, D. 1976. Unpublished term paper, UCLA TESL Department.

Kessler, C. 1975. Postsemantic processes in delayed child language related to first and second language learning. In D. Dato (ed.), *Developmental psycholinguistics: theory and applications*, Georgetown University Roundtable on Languages and Linguistics 1975. Washington, D.C., Georgetown University Press. Pp. 159-178.

Kessler, C., and I. Idar. 1977. The acquisition of English syntactic structures by a Vietnamese child. Paper presented at the Los Angeles Second Language Research Forum, UCLA. February, 1977.

Krashen, S. 1976. Formal and informal linguistic environments in language learning and language acquisition. *TESOL Quarterly 10*, 157–168.

Krashen, S. 1977. The Monitor Model for adult second language performance. In M. Burt, H. Dulay, and M. Finocchiaro (eds.), *Viewpoints on English as a second language*, New York, Regents. Pp. 152–161.

Krashen, S. Individual variation in the use of the Monitor. In W. Ritchie (ed.), *Second Language Acquisition Research: Issues and Implications*, New York, Academic Press. In press.

Krashen, S. The monitor model of second language performance. Papers presented at the 1976 California Linguistics Association Conference. In press.

Krashen, S., P. Pon. 1975. An error analysis of an advanced ESL learner: The importance of the Monitor. *Working Papers on Bilingualism* 7, 125–129.

Krashen, S., and H. Seliger, 1976. The essential contributions of formal instruction in adult second language learning. *TESOL Quarterly* 9, 173–183.

Krashen, S., J. Butler, R. Birnbaum, and J. Robertson. 1976. Two studies in language acquisition and language learning. Paper presented at the Linguistic Society of America, Philadelphia. December, 1976. Also to appear in the *CUNY Forum*.

Krashen, S., N. Houck, P. Giunchi, S. Bode, R. Birnbaum, and J. Strei. Difficulty order for grammatical morphemes for adult second language performers using free speech. *TESOL Quarterly 11*, 3.

Lackstrom, J., L. Selinker, and L. Trimble. 1973. Technical rhetorical principles and grammatical choice. *TESOL Quarterly* 7, 127–136.

Larsen-Freeman, D. 1975. The acquisition of grammatical morphemes by adult learners of English as a second language. Doctoral Dissertation, University of Michigan.

Lawler, J., and L. Selinker. 1971. On paradoxes, rules, and research in second language learning. *Language Learning 21*, 27–43.

Linde, C., and W. Labov. 1975. Spatial networks as a site for the study of language and thought. *Language 51*, 924–939.

LoCoco, V. 1975. An analysis of Spanish and German learner's errors. *Working Papers on Bilingualism* 7, 96–124.

Newmark, L. 1966. How not to interfere with language learning. In *Language Learning: The Individual and the Process. International Review of American Linguistics 40*, 77–83.

Plann, S., and A. Ramirez. 1976. A study of the English and Spanish of Spanish speaking pupils in a Spanish immersion school program. *UCLA Working Papers in TESL*, 1976. Pp. 89–108.

Politzer, R. 1968. An experiment in the presentation of parallel and contrasting structures. *Language Learning 18*, 35–53.

Postovsky, V. 1977. Why not start speaking earlier? In M. Burt, H. Dulay, and M. Finocchiaro (eds.), *Viewpoints on English as a second language*, New York, Regents. Pp. 17-26.

Richards, J. 1975. The context for error analysis. In M. Burt and H. Dulay (eds.), *New directions in second language learning, teaching, and bilingual education: On TESOL '75*, Washington, TESOL. Pp. 70–79.

Rosansky, E. 1976. Methods and morphemes in second language acquisition research. *Language Learning 26*, 409-425.

Schlue, K. 1976. An inside view of interlanguage. MA Thesis. UCLA TESL Department.

Selinker, L., M. Swain, and G. Dumas. 1975. The interlanguage hypothesis extended to children. *Language Learning 25*, 139–155.

Snow, C., and M. Hoefnagel-Hohle. Age differences in second language acquisition. In E. Hatch (ed.) *Second Language Acquisition*. Rowley: Newbury House. In press.

Stafford, C., and G. Covitt. Monitor use in adult second language production. *ITL: Review of Applied Linguistics*. In press.

Strei, G. 1976. Acquisition data project. Term Paper, Department of Linguistics, USC.

Taylor, B. 1975. The use of overgeneralization and transfer learning strategies by elementary and intermediate students in ESL. *Language Learning 25*, 73–107.

Wagner-Gough, J., and E. Hatch. 1975. The importance of input data in second language acquisition studies. *Language Learning 25*, 297–308.

Interference versus Structural Complexity in Second Language Acquisition: Language Universals as a Basis for Natural Sequencing

Georgette Ioup and Anna Kruse
University of Washington

A current controversy in the field of second language acquisition centers around the issue of whether interference or intralingual factors most affect second language acquisition. The present study examines the factors responsible for errors by investigating second language acquisition of a particular English structure—the relative clause. In an attempt to define a potential natural sequencing for this structure, we refer to the recent linguistic research on language universals. Three linguistic hypotheses make predictions as to which relative clause structures will be easiest for first language learners to acquire. Kuno (1974) argues that non-center embedded structures will present less difficulty to the learner than those which are center embedded. Keenan (1972) theorizes that relative clauses with relativized subjects will be acquired before those with relativized objects. Sheldon (1973) claims that relative clauses exhibiting a parallel function with the head noun will be acquired more easily than those which do not. Though each of the hypotheses makes different predictions concerning which relative clause structures will present most difficulty to the learner, all three positions are supported with data from first language acquisition. In order to determine which hypothesis makes correct predictions for second language learning, data is collected from adult learners from various language backgrounds. Results indicate that interference plays a minimal role and that the language universals defined by both Keenan and Kuno offer insights into the linguistic behavior of our subjects. These findings, while contributing to our knowledge of the particular processes involved language development, also have important implications for the general theory of language acquisition.

Two questions are very central in current studies of adult second language acquisition. Researchers are asking to what extent a learner's competence at any given stage is influenced by knowledge of the native language, and the related question of how closely the adult's acquisitional stages approximate first and second language acquisition in children. Hypotheses addressing these questions were first put forward by Fries (1945) Lado (1957) and their contemporaries. They claimed that all difficulty in learning a second language arose from the interference effects of the native language and not from any inherent difficulties in the target language. Fries, in a forward to Lado (1957), unambiguously states this position:

> Learning a second language, therefore constitutes a very difficult task from learning the first language. The basic problems arise not out of any essential difficulty in the features of the new language themselves but primarily out of the special "set" created by the first language habits.

Such a narrow view of second language acquisition was soon challenged. Many researchers (for example Pit Corder, 1971; Nemser, 1971; Richards, 1971; and Selinker, 1972) insisted that a variety of factors were responsible for the problems encountered in learning a second language. Intrinsic intralingual complexities were claimed to influence a learner's developmental progress to as great an extent as native language transfer. Richards (1971) is clear on the importance of both learning variables:

> Interference from the mother tongue is clearly a major source of difficulty in second language learning, and contrastive analysis has proved valuable in locating areas of interlanguage interference. Many errors however, derive from the strategies employed by the learner in language acquisition, and from the mutual interference of items within the target language. (p. 214)

Empirical work remained to be done to substantiate the newer contentions. As studies attempted to confirm the role of intralinguistic factors, the importance of interference was downgraded. A new account of second language learning was then proposed which excluded native language transfer as a significant learning variable. It's chief advocates were Dulay and Burt (1974a & b). Their position gained credence as a result of several studies which demonstrated little evidence of interference in both child and adult L_2 acquisition. One such study, researched by Ervin-Tripp (1974), investigates the learning experience of English speaking children acquiring French. Interference could not account for their linguistic behavior in both elicited and spontaneous productions, leading Ervin-Tripp to conclude that "second language learners recapitulate mother tongue acquisition" (p. 121). Dulay and Burt (1974a) come to similar conclusions after studying the acquisition of English morphology by Chinese and Spanish speaking children. Morpheme difficulty orderings for the two groups were established and were shown to be highly correlative, findings which argue for the L_2 system, rather than the L_1 system, as a guide to L_2 acquisition. This leads them to reject first language habits as productive predictors of children's acquisition of second language syntactic processes.

Other researchers have extended these results to adult ESL learners. Cook (1973) discovers that child first language learners and adult second language learners produce the same types of errors in imitation tasks involving English relative clause structures. Bailey, Madden and Krashen (1974) replicate the Dulay and Burt morpheme acquisition study with adult learners and determine that a similar difficulty ordering obtains between both the child and adult L_2 learners.

The issue, however, is far from resolved. Though these studies do not attest significant native language interference, other studies do. Hakuta (1975) in a case study of a Japanese child learning English observes errors in her spontaneous speech which he can only attribute to transfer. His data include cases of lexical transfer and evidence for structural avoidance. He rejects the thesis advocated by Dulay and Burt and recommends in its place a return to the interlanguage hypothesis which accounts for the learner's stages of competence through an interaction of three factors; overgeneralization, simplification and transfer.

In two different studies Schachter (1974) and Schachter, Tyson and Diffley (1976) report evidence of transfer in adult acquisition. Their studies are noteworthy because they systematically compare subjects from several different language backgrounds along dimensions which are selected because of the particular difficulties they are expected to present to each language group. If it is the case that the structures are either randomly troublesome, or difficult for all (or none) of the language groups, it becomes necessary to reject the interference hypothesis for these structures. It would then be necessary to attribute any difficulties to the complexities inherent in the structures. Conversely, should the predictions be borne out, the interference theory would be confirmed.

Unfortunately the results of both studies are mixed and do not provide clear support for either position. In the first study Schachter examines only spontaneous performance data. Some expected errors are observed, others are not. Schachter concludes that avoidance of structural types is evidence that the structure is posing problems for that particular language group. We contend that there is no justification for this conclusion. One never knows why speakers choose to express their thoughts using one syntax rather than another. If mother tongue interference is playing a role it may be at the sylistic level rather than the structural level. Avoidance gives us no clues as to why a learner might find that structural type difficult. Studying performance, therefore, can only provide limited insight into the learner's level of competence at a particular stage of acquisition.

Schachter, Tyson and Diffley acknowledge the limitations of performance studies in designing a second investigation. They argue convincingly that it is necessary to gather intuitional data in order to completely characterize a learner's knowledge at some particular stage. A description of the learner's system which is based entirely on spontaneous production data will account for only a small fragment of the learner's knowledge of the target language. This study elicits grammaticality judgments on English sentences from speakers with five different native language backgrounds: Japanese, Chinese, Persian, Arabic and Spanish. Unfortunately, the results again are mixed and no firm conclusions as to the role of interference can be drawn from their investigation.

So despite an abundance of current research, there are still no firm answers to the questions posed at the beginning of this paper. Researchers disagree on the extent to which transfer influences second language acquisition and the data present conflicting results. But until we are able to define the role of interference precisely, it will be impossible to isolate a set of intralinguistic variables and assess their role in learning. Thus it is necessary to conduct research designed to separate interference variables from intralinguistic variables and then determine the relative difficulty of structures in each group. In doing this, we not only shed light on the processes involved in second language learning, but also contribute to our knowledge of language universals and the general principles of language acquisition.

The research we are reporting on focuses on the acquisition of the English relative clause. In the first portion of the study we investigate the role of interference; in the second part we attempt to define a natural difficulty ordering for the stages of relative clause acquisition. The research incorporates two of the operational principles outlined in Schachter (1974) and Schachter et al. (1976). Subjects are selected from a number of very different language backgrounds. The native languages involved are Spanish, Persian, Japanese, Arabic and Chinese. Intuitional judgments rather than spontaneous or elicited productions provide the data base for our investigation.

Research Design

Relative clause structures from our subjects' native languages were compared to English along four dimensions, shown in Table 1. A contrastive analysis (CA) theory of interference as defined in Lado (1957) and advocated in Schachter (1974) assumes that we can predict which aspects of a structure will cause difficulty by determining where the points of structural contrast lie. Any distinct difference is expected to impede the acquisition of the structure along that dimension. CA further argues that a partial overlap, where one unit in the native language corresponds to two or more units in the target language will present maximum learning difficulty. Table 1 reflects these assumptions by underlining the cells in each column where partial overlap occurs and therefore where the CA hypothesis anticipates the greatest number of errors.

TABLE 1

Classification of Relative Clauses Along Four Dimensions

	Pronoun Retention	R.C. Position	Rel Marker Appearance	Rel Marker Morphology
Chinese	*sometimes*	*precedes head*	always	invariable
Japanese	none	*precedes head*	none	———
Arabic	*sometimes*	follows head	*sometimes*	invariable
Persian	*sometimes*	follows head	always	invariable
Spanish	none	follows head	always	*variable*
English	none	follows head	sometimes	variable

We observe then that in terms of pronoun retention inside the relative clause, Chinese, Arabic and Persian exhibit partial overlap with English, retaining the pronoun in some NP positions but not in others; while Japanese and Spanish pattern like English. Thus we expect high levels of difficulty from Chinese, Arabic and Persian speakers within the category of pronoun retention, but little or no difficulty for Japanese and Spanish speakers.

Relative clauses may either precede their head noun as in Chinese and Japanese; or follow the head noun as in Arabic, Persian, Spanish and English. We would predict then, that Japanese and Chinese speakers would exhibit the most errors here.

All of the languages studied except Japanese use a relative marker with the relative clause (e.g. English *that* or *who*). In English, the marker is often deleted in the surface form of the sentence by a rather complex rule. It states, in general, that a relative pronoun may be optionally deleted if the embedded verb will not then be mistaken as a main verb. Arabic is the only other language in our group which permits deletion of the relative marker. In Arabic the rule for deletion is quite different from the English rule and it functions obligatorily. The relative marker must be used when the antecedent is definite, but is obligatorily omitted with indefinite antecedents. The CA hypothesis predicts that though speakers from all five language backgrounds will have difficulty mastering the English rule, the Arabic speakers will have greatest difficulty.

Our fourth category is relative marker morphology. In Spanish and English the markers are inflected, but according to different inflectional criteria. The other languages which have markers do not inflect them at least not in their spoken forms. Therefore, although all five language groups are expected to have difficulty mastering the English inflectional system, we anticipate more errors among Spanish speakers due to the partial overlap between the two languages.

Method

We elicited grammaticality judgments by presenting subjects with a written list of 36 sentences each containing a relative clause structure. Subjects were asked to judge them as structurally correct or incorrect. The malformed relative clauses varied along the dimensions specified in Table 1 and the well-formed sentences were those to be used in the second part of this study, and will be referred to later in the paper. The sentences were presented in random order. Examples of the malformed relative clauses are listed in Table 2. (For a complete list of the deviant sentences used in this study, refer to the Appendix.)

TABLE 2

Sample Deviant Sentences

1) Pronoun retention in the relativized position
   *The bed which the boy put the shoes under *it* is in the corner.
2) Relative Clause Preposing
   *The eats a big breakfast girl is very strong.
3) Relative Marker Deletion
   *The girl was last in line didn't get a ticket.
   The woman James met is my sister. (Correct in English)
4) Incorrect Pronoun Morphology
   *John hates which his brother likes.

The subjects consisted of 18 speakers of Spanish, 16 speakers of Chinese, 18 speakers of Persian, 10 speakers of Japanese and 25 speakers of the Saudi Arabian dialect of Arabic. These 87 subjects were randomly chosen from among low and middle intermediate level students enrolled in English as a Second Language classes at four Seattle area universities. Subjects were instructed to read the sentences and to place an 'X' in front of any that they felt were not correct sentences of English. They were told to concentrate on the grammar and to ignore strangeness in meaning or questions of spelling and punctuation. There was no time limit.

Results

The main question we are attempting to deal with at this point is whether it is language group, sentence type, or a combination of the two that is more predictive of a learner's performance. In order for our results to establish a transfer effect, it is necessary that either the main effect of language group or the interaction of language group with particular sentence types be significant as compared with the variation across individual subjects. Table 3 indicates the mean score of incorrect responses per cell.

TABLE 3

Analysis of Deviant Sentences: Mean Score of Incorrect Responses

	Pronoun Retention N = 8	R.C. Position N = 4	Rel Marker Appearance N = 6	Rel Marker Morphology N = 8
Chinese N = 16	49.2	23.4	47.9	24.2
Japanese N = 10	35	25	38.3	22.5
Arabic N = 25	68.5	31	40.7	47.5
Persian N = 18	37	12.5	37	11.1
Spanish N = 18	47.2	27.8	41.7	45.8

To evaluate the data, we used a two way analysis of variance design modified to screen out unequal N's, repeated measures and nested subjects, so that we could determine whether any group response differed significantly from the individual variation within groups. We were interested to see initially whether language group, sentence type, or a combination of the two predicted performance to any extent. Significance was set at the .001 level since an effect was predicted. Our findings show that the main effect of language group did not turn out to be significant, neither did the interaction between language group and sentence type. However, we found the main effect of sentence type to be highly significant. These results indicate that although the main scores for each language group on a particular sentence type *look* very different, the variation among the individual scores was great enough that the differences were not actually significant. Thus any differences between mean scores for language groups on a given sentence type might well disappear in testing further samples. Since the interaction of language group and sentence type was not found to be significant, we were assured that no pattern of ease for one language group and difficulty for another on a given sentence type was cancelling out a main effect. Therefore, we were assured that the only effect that was significant was that of language type. The lack of any significant interaction precluded the necessity for finer statistical analysis.

Our results show that, contrary to the CA hypothesis, sentence type rather than native language background is the most reliable predictor of error. The question then arises as to why certain of the structural parameters tested are more difficult to master than others. In order to answer this question we must first determine whether the four sentential categories can be arranged into a difficulty ordering. To do this, we totaled the mean scores of all subjects on a sentence type—having ruled out the individual language factor—and did a series of difference of means t-tests, adjusting for the same error term as required for the analysis of variance. Significance was set at the .01 level for a one-tailed test. The total mean score for each sentence type is presented in Table 4. The differences were found to be significantly different enough to permit them to be ordered.

TABLE 4

Total Mean Score of Incorrect Responses According to Sentence Type

Pronoun Retention	R.C. Position	Rel Marker Appearance	Rel Marker Morphology
47.38	23.94	41.12	30.22

When we rank order the four structural categories in terms of their relative difficulty, we arrive at the following difficulty ordering. Decisions concerning pronoun retention are the most problematic, followed by questions of relative marker deletion, relative marker morphology and relative clause position, in that order. We can explain the low error total on relative marker position since fundamental word order is one of the earliest concepts acquired in learning a new language, both first and second. Relative pronoun morphology will cause some learning difficulty, but, in learning to form the relative clause, the rules governing morphological distribution are the ones given particular attention in the classroom and are usually practiced and tested until error free performance is reached. The rule for marker deletion is complex and optional in English, therefore it is not surprising that our subjects who are not very advanced have not yet mastered it. What is surprising is that pronoun retention is a problematic area for all language groups, even those where such errors are not expected.

Furthermore, in talking with our subjects' language teachers we found that problems of pronoun retention receive special focus in the classroom. How then can we explain the high occurrence of errors in this category for all language groups?

E. L. Keenan (1972a) argues that pronominal reflexes at the surface level are common in many languages of the world precisely because they are very natural manifestations of the underlying logical structure of the language. Furthermore, they are retained at the surface most frequently when an aid to comprehension is required due to the structural complexity of the sentence. These pronouns function cognitively to assist in perceptual coding. We propose that language learners systematically employ pronominal reflexes in the beginning stages as a perceptual heuristic to simplify the coding task. In fact, we have noticed that advanced learners who do not ordinarily do so will employ pronominal reflexes quite naturally when sentences reach a certain level of complexity, as for example: "That's the man who Sally told Ellen her mother was talking about him."

Keenan (1972b) has further theorized that there exists a natural ordering within the category of pronoun retention, with genitive and preposition object pronouns the most likely to be retained, and the retention of subject and direct object pronouns the least common. The ordering is based on the observation that in natural language, relative clause formation is more natural and hence more accessible with relative clause subjects and objects than with genitives and preposition objects. He has demonstrated that sentences containing relative clause structures formed on these latter two positions are perceived as structurally more complex than corresponding sentences with the relative clause formed on either the subject or direct object. It would follow then that our subjects would accept more readily a pronominal reflex in the less accessible relative clause positions.

Since the study included sentences exhibiting all tour types of pronoun retention, we totaled the scores across language for each category and compared them to ascertain if subjects did indeed find it more natural to retain a pronoun in the less accessible relative clause positions. The test permitted 174 possibilities for error in each category. Total error count for each grammatical category is presented in Table 5.

TABLE 5

Total Error Count for Categories of Pronominal Reflex

Subject	Direct Object	Prepositional Object	Genitive
64	64	106	104

We observe that the genitive and preposition object positions produced approximately one third more errors than either the subject or direct object positions. A difference of proportion t-test was done to determine if these differences were significant. Significance was found at the .01 level indicating that the natural difficulty ordering for relative clause formation proposed by Keenan provides a reasonable explanation of our subjects behavior.

Testing for Natural Sequences

Since the first part of the study was not able to locate significant structural interference in the acquisition of the English relative clause, an attempt was made to locate other factors which might account for the learning difficulties. The second part of this study attempts to ascertain if there exists a natural difficulty ordering among the various types of relative clause structures, resulting from the strategies learners are using to process them. If such an ordering can be shown to exist, we then want to determine to what extent it reflects the acquisition strategies employed by first language learners of English. As we searched the linguistic literature for possible hypotheses to test, we were surprised to discover that there were three such hypotheses each making conflicting predictions about the same set of sentences. These sentences are listed below in Table 6.

TABLE 6

Sentences Used to Establish a Difficulty Ordering

SS	The dish which fell on the floor broke in half.
	The picture which won the prize is hanging on the wall.
SO	The sweater which I found on the bus belongs to Susie.
	The candy which Billy gave me tasted good.
OS	The little girl is looking for the cat which ran away.
	The old woman picked some flowers which were growing in the park.
OO	The boys are reading the books which they borrowed from the library.
	The child ate the cookies which the neighbors baked.

Kuno (1974) argues that center embedding reduces comprehensibility causing relative clause structures which are center-embedded to be perceptually more difficult than right branching clauses. The perceptual problems of nested constructions resulting from a strain on memory capacity were first predicted by Chomsky and Miller (1963). In relation to first language acquisition, Slobin (1971) predicts the late acquisition of center-embedded clauses, based on his operating principle D which states that there is pressure to avoid interruption of underlying linguistic units. Menyuk (1969) observes that sentence final relative clauses appear first in the data before center-embedded ones. In our study this hypothesis predicts that the sentence types we have labeled OS and OO will be easier to process.

A second hypothesis is developed by E. L. Keenan (1975). As mentioned earlier Keenan has established an accessibility hierarchy for relative clause structures where those relative clauses with relativized subjects occur higher on the hierarchy than those with relativized objects. He predicts that the more accessible relative clauses will be perceived as more natural and hence will be easier to process and learn. In testing done with school age children, he found that retention was better on relative clause structures higher on the hierarchy. In English, Keenan's hypothesis receives independent support from another of Slobin's principles, Universal D. It states that structures requiring permutation of elements will first appear in non-permuted form. Menyuk (1969) has observed that the earliest relative clauses do indeed appear without inversion. Keenan's prediction, then, is that the SS and OS sentence types will be considered grammatical more often than the SO and OO types.

A third hypothesis is put forth by Sheldon. Through testing she has done, Sheldon has ascertained that children have better comprehension of those relative clause structures where the relative pronoun performs the same function as the head. The parallel function hypothesis is presented to account for the processing strategies children use to interpret relative clauses. It follows from her hypothesis that relative clauses manifesting a parallel function will be acquired before those that do not. In reference to this study it predicts that the SS and OO sentences should be the least difficult to process. These three hypotheses are summarized in Table 7.

TABLE 7

Summary of Hypotheses Tested

1. Kuno:			Center Embedding is perceptually difficult and should be avoided. Supported by Slobin's Operating Principle D: Avoid interruptions of underlying linguistic units.
	OS	SS	
	OO	SO	
2. Keenan:			The Accessibility Hierarchy will predict the order of acquisition: relative clauses formed on subjects will be easier than those formed on subjects. Supported by Slobin's Universal D: Avoid rearrangement of underlying word order.
	SS	SO	
	OS	OO	
3. Sheldon:			The Parallel Function Hypothesis: relative clauses where the relative pronoun has the same function as the head will be easier to acquire.
	SS	SO	
	OO	OS	

Results

The three hypotheses make predictions concerning the same set of sentences, however, the predictions are conflicting. Each theory predicts that a different pair of relative clause structures will be easier to process and acquire. The total number of errors for each sentence type is given in Table 8. Each category permits 174 possibilities for error. It is evident that subjects recognized the majority of well-formed relative clauses as grammatical.

TABLE 8

Total Errors for Each Sentence Type

SS	45 (39%)
SO	30 (26%)
OS	13 (11%)
OO	26 (23%)

The statistical analysis used in the second part of the study was again a two way analysis of variants, adjusted for nesting, repeated measures and unequal cases per cell. The analysis yielded the following results. Errors according to language group were not significant at the .05 level, while errors according to sentence type were significant. However, there was no significant interaction.

Errors were then totaled according to hypothesis. These figures are presented in Table 9. A modified t-test was used to determine the significance of the difficulty ordering obtaining between the pairs as specified in the three hypotheses. The Kuno hypothesis was confirmed, the Keenan hypothesis was not. Not only was the Sheldon hypothesis not confirmed, but the results go significantly in the opposite direction of what is predicted.

TABLE 9

Total Errors by Hypothesis

	Easier	Harder
Kuno	OS OO = 39 (32%)	SS SO = 75 (68%)
Keenan	SS OS = 58 (51%)	SO OO = 56 (49%)
Sheldon	SS OO = 71 (62%)	SO OS = 43 (38%)

Our data therefore supports only the Kuno hypothesis establishing center embedding as a parameter of perceptual complexity in relative clause processing. However, we do not believe that our results speak conclusively. Additional research needs to be conducted on this issue. Our testing procedures allow for reflection on the part of our subjects and as such would be classified by Krashen (1976) as "slow testing". Therefore, we are measuring both the acquired and learned competence of our subjects, as we have no way of factoring out one of the measures. But it is necessary that a natural difficulty ordering be based on acquired knowledge alone, if it is to provide a comparison of first and second language acquisition.

Conclusion

Our research was specifically designed to test the transfer hypothesis specified by Lado, and it was found to be wanting. With respect to the syntactic aspects of language, Lado confines his analysis to word and surface morphology. By and large, the studies which have not demonstrated a transfer effect have restricted their investigations to these structural parameters, while other studies which claim evidence of interference have extended the scope of their research to include semantic, pragmatic and stylistic variables. Perhaps Lado's variables are not the aspects of the native language which interfere with the acquisition of a new language. Yet Lado's account of interference is the only lucid treatment that exists to date. No linguist since Lado has attempted a systematic analysis of the linguistic variables in the native language which influence hypothesis formation in learning a new language. Furthermore, to claim that transfer is a viable factor in second language learning, is it necessary to demonstrate the intrusion of native language structures and strategies at every point where a contrast between the two languages occurs? If not, then we require a new hypothesis defining precisely where transfer is expected and an explanation as to why it is confined to these aspects of language acquisition. Until such a theory is developed to guide our research, we can expect that studies investigating different components of second language acquisition will continue to come to radically different conclusions about the role of L_1 competence in acquiring L_2.

REFERENCES

Bailey, N. C. Madden, and S. Krashen. 1974. Is there a "natural sequence" in adult second language learning? *Language Learning 24*, 2.

Chomsky, N., and G. Miller. 1963. Finitary models of language users. In Luce, Bush and Galanter (eds.), *Handbook of mathematical psychology* II, New York, John Wiley and Sons.

Cook, V. J. 1973. The comparison of language development in native children and foreign adults. *International Review of Applied Linguistics 11*, 1.

Corder, S. P. 1971. Idiosyncratic dialects and error analysis. *International Review of Applied Linguistics* 9, 2.

Dulay, H., and M. Burt. 1974a. Natural sequences in child second language acquisition. *Language Learning 24*, 1.

Dulay, H., and M. Burt. 1974b. Errors and strategies in child second language acquisition. *TESOL Quarterly 8*, 2.

Ervin-Tripp, S. 1974. Is second language learning like the first? *TESOL Quarterly* 8, 2.

Fries, C. 1945. *Teaching and learning English as a foreign language*. Ann Arbor, University of Michigan Press.

Hakuta, K. 1975. *On becoming bilingual at age five: the story of uguisi*. B.A. Thesis. Harvard University.

Keenan, E. L. 1972a. On semantically based grammar. *Linguistic Inquiry* 3, 4.

Keenan, E. L. 1972b. The logical status of deep structures. Paper presented at the 11th International Congress of Linguists, Bologna, Italy.

Keenan, E. L. 1975. Variation in universal grammar. In R. Fasold and R. Shuy (eds.), *Analyzing Variation in Language*, Washington, Georgetown University Press.

Krashen, S. 1976. Formal and informal linguistic environments in language acquisition and language learning. *TESOL Quarterly 10*, 2.

Kuno, S. 1974. The position of relative clauses and conjunctions. *Linguistic Inquiry 5*, 1.

Lado, R. 1957. *Linguistics across cultures*. Ann Arbor, University of Michigan Press.

Menyuk, P. 1969. *Sentences Children Use*. Cambridge, Mass., MIT Press.

Nemser, W. 1971. Approximative systems of foreign language learners. *International Review of Applied Linguistics* 9, 2.

Richards, J. 1971. A non-contrastive approach to error analysis. *English Language Teaching* 25, 3.

Schachter, J. 1974. An error in error analysis. *Language Learning* 24, 2.

Schachter, J., A. Tyson, and F. Diffley. 1976. Learner intuitions of grammaticality. *Language Learning* 26, 2.

Selinker, L. 1972. Interlanguage *International Review of Applied Linguistics* 10, 3.

Sheldon, A. 1973. The role of parallel function in the acquisition of relative clauses in English. Mimeograph. Indiana University Linguistics Club.

Slobin, D. 1971. Developmental psycholinguistics. In W. Dingwall (ed.), *A survey of linguistic science*, College Park, University of Maryland, Department of Linguistics.

APPENDIX

Malformed Test Sentences

Pronoun Retention

Subj:	The boy saw the girl who *she* hit him. The girl who *she* chased the bird fell down.
Obj:	The baby which the man is holding *it* is crying. The woman saw the man whom she photographed *him*.
Prep Obj:	The girl whom the dog barked at *her* fell down. The bed which the boy put the shoes under *it* is in the corner.
Poss:	We found the baby *that his* mother went away. The boy *that* the dog bit *his* hand is crying.

Relative Clause Preposing

Subj:	I am looking for the in my purse comb. The eats a big breakfast girl is very strong.
Obj:	The Susan wrote a letter to girl never answered it. The milk spilled on the Martin just washed floor.

Relative Marker Deletion

Subj:	The man liked the baby was smiling. The girl was last in line didn't get a ticket.
Obj:	The woman James met is my sister. I ate the food Fred cooked. (Correct in English)
Poss:	The child's mother hit him is crying. I helped the boy's father is sick.

Relative Marker Morphology

who:	I like people which are friendly. The girl which discovered the money received a reward.
which:	The book whom you're looking for is on the table. John threw the ball who broke the window.
whose:	The man which hat I found in the park visited me. My brother knows the boy which are was broken.
what:	I kept who Susan gave me. John hates which his brother likes.

A Rationale for Discourse Analysis in Second Language Acquisition Research

Diane Larsen-Freeman

University of California, Los Angeles

This paper introduces discourse analysis as an interesting and potentially fruitful approach to studying second language acquisition. Although discourse analysis can include a study of both written and spoken texts, the main thrust of discourse analysis in L_2 acquisition has dealt with the semantic and communicative functions of spoken utterances, the structure of the spoken language at the suprasentential level and the interaction of native-speaker speech speech input with the learner's speech product.

Interesting questions being addressed through application of this new approach include:

1. What is the structure of language at the discourse level?
2. How does L_2 acquisition take place through conversational interaction?
3. What can conversational analysis contribute to the optimal age issue?

Until recently, researchers in second language acquisition have focused their attention on the speech product of the learner. Transcripts of the learner's speech have been examined for evidence of the acquisition of grammatical rules which govern the form of particular structures such as relative clauses, plural markers, propositional negation, etc. The criterion for acquisition has varied, but basically has depended on the learner's use of a particular structure in an obligatory context.

From observation on the use of a structure by the learner, researchers would then infer certain strategies adopted by the learner in the second language acquisition process. Strategies such as creative construction (Dulay and Burt, 1973), generalization (Richards, 1971), language transfer (Selinker, 1972) and others have been proposed.

While studies conducted employing this approach have made important contributions to our knowledge of second language acquisition, researchers have begun to wonder if by focusing their attention on the use of grammatical forms in the learner's speech, their perspective might not be too limited.

Wagner-Gough (1975) has pointed out that a learner's production of a particular structure alone might be an inadequate measure of the learner's acquisition of that structure. She argues that an assessment of his ability to handle all the functions of the structure should be undertaken as well. Just because the learner attaches an "ing" to the main verb following the "be" auxiliary to produce the present progressive tense, for example, does not necessarily imply that the learner knows he can use this same form to express an action occurring right now as well as one that will occur in the future.

Not only has the semantic function of a structure often been overlooked, but the communicative function of a structure has as well. For instance, if a learner produces an utterance such as "How are you?" in an appropriate context, he is not only perhaps demonstrating his knowledge of the syntactic form of WH-Q's and his knowledge that such a form has the semantic function of requesting information, but also his knowledge that this is a question often given the task of opening or initiating conversation.

Indeed, in focusing only upon the structures at the sentential level, we have perpetrated a misleading simplification of the language acquisition process. We have overlooked the need for the learner to acquire a whole other system of language—namely the structural unity that exists at the discourse level. There is a rule-bound system which exists at the suprasentential level in both conversation (Sacks, Schegloff and Jefferson, 1974) and in written discourse (Halliday and Hassan, 1976) that the learner must also master if he is truly to acquire a second language.

In addition, it has been suggested that by focusing solely on the linguistic form of the learner's speech product, we have virtually ignored an important data source, namely the language input to which the learner is exposed. Since an examination of the input is important for pointing out additional acquisition strategies the learner is adopting such as incorporation of formulas (Hatch and Wagner-Gough, 1974) and imitation or matching (Larsen-Freeman, 1976), the input must be considered as well.

Some researchers would go even farther and say that it is the study of the interaction of the input and the linguistic product which is most enlightening. Indeed, in a first language acquisition study, Scollon (1974) advances the notion that a construction within a child's speech may at first develop out of a construction that the child builds jointly with other speakers. Scollon calls each such exchange a vertical construction. For example:

Child: Kimby.
Mother: What about Kimby?
Child: close (i.e., "closed")

Although the child's message is obscure, Scollon proposes that since the adult has taken the child's first utterance as the statement of a topic and has responded by asking the child to comment, it might be from interactions like these that a child learns how to produce topic-comment constructions.

Hatch (1976) suggests this same process might be operating in second language acquisition. Rather than a learner acquiring basic syntactic patterns and then learning how to invoke them properly in discourse, Hatch wonders whether a more accurate portrayal might be that the "learner learns how to converse, how to interact verbally and out of these interactions syntactic structures are developed."

Thus, a methodology which looks at the semantic and communicative functions of a structure, the structural unity at the suprasentential level, the input to the learner and perhaps the input/product interaction might yield some interesting insights that we have been denied by focusing solely on the forms in the learner's speech product. A methodology which purports to do just this is discourse analysis.

While not all researchers have embraced this new methodology, some are beginning to tackle important questions about second language acquisition using a discourse analysis approach.

Some of the interesting questions, as I perceive them, being addressed through application of this new approach are outlined below.

What is the Structure of Language at the Discourse Level?

Although certain linguists, particularly of the British school (notably Firth and Halliday) have been interested in this area for a long time, researchers from other disciplines have begun to make a contribution. Indeed, anthropologists, Sacks, Schegloff and Jefferson (1974), have inspired a whole field now referred to as conversational analysis. Using the conversational analysis framework, second language acquisition researchers, Schwartz (1977) and Gaskill (1977) have been examining ESL learners' conversations to see if the observations made by Schegloff, Jefferson and Sacks (1977) about native speaker conversation hold for second language learner conversation as well.

While Schwartz' study of conversations between pairs of non-native speakers and Gaskill's study of the conversational interaction between native and non-native speakers focus primarily on the self and other-corrections made by speakers in conversations, Keller-Cohen (1977) has been investigating the different ways child non-native speakers attempt turn allocations when engaged in conversation with adult native speakers.

Dealing with conversational data from a different perspective, Arthur (1977) has conducted a study which investigates the type of termination-of-conversation strategies employed by native and non-native speakers. He was also interested in seeing if the sex of the speaker influenced the termination strategy employed.

Godfrey (1977) has also been attempting to analyze the structure of language at the discourse level. He has, however, chosen to examine the discourse features present in non-native speaker monologues in English, as opposed to the interactional component present in conversation.

Vanderbrook, Schlue and Campbell (1977) and Celce-Murcia (1977) have been employing a "textual analysis" to determine the way context, a speaker's presupposition and the discourse function of a structure influence the form of the structure in discourse. For example, Vanderbrook, Schlue and Campbell discovered certain conditions when it was likely for native speakers to produce uninverted yes/no questions in discourse.

From the brief survey of some of the studies being pursued under the discourse analysis rubric, we can begin to appreciate the variety of ways this first question is being addressed.

How Does Language Acquisition Take Place Through Conversational Analysis?

In a first language acquisition study, Cosaro (undated manuscript) found a great deal of language teaching occurring in adult to child speech. Cosaro maintains that the adult, consciously or unconsciously, teaches the child about language and its use through discourse. In Cosaro's data we find examples of the adult guiding a child through conversation by supplying the child with both a question and an appropriate answer within a single turn:

Adult: What's that Mia's got, a horse?
Child: Horse!

It is not impossible that we could find a similar kind of question and answer interaction pattern in teacher speech to learners. From this sort of input the learner would be able to induce what he was being asked and be given a relevant reply with which to respond.

In both first and second language acquisition studies (Scollon, 1974; Hakuta, 1975; Hatch and Wagner-Gough, 1974; and Fillmore, 1976) we find reports that much of the speech of the learner consists of prefabricated routines—patterns or formula from native speaker speech which the learner incorporates into his repertoire without first analyzing. An example that we might find second language learners adopting from teacher speech input would be the pattern: This is an *NP*. If this pattern is frequently occurring in the discourse input to the learner, it is conceivable that the learner would adopt and use it holistically without ever stopping to analyze the constituents rules comprising the pattern.

Scollon's vertical constructions, Cosaro's question and answer interaction and incorporation are but a few of the possible means by which a learner discovers how to produce acceptable utterances from exposure to discourse. No doubt this particular question will be addressed with increasing vigor as the field evolves.

What Can Conversational Analysis Contribute to the Optimal Age Issue?

At one time or another second language acquisition researchers have entertained the thought that one, all, none or a combination of the following could be used to explain the purported differential success between child and adult learners of a second language: biological factors, affective factors, motivation, time allotment, cerebral dominance (hemisphericity) and learning conditions.

Discourse analysis, too, can possibly shed some light in this area—or at least propose another hypothesis. After surveying data collected by a number of researchers, Hatch (1976) offers the opinion that the input differences to children and adults vary considerably. While topics nominated to adult nonnative speakers are varied and abstract, an adult native speaker does not suggest topics to a child second language learner for which the child has no background information. The topics discussed by the adult and child are very much focused in the here and now for which there are ample cues from context for the child to discern meaning. Thus, the task of language learning through discourse is presumably simpler for the child than the adult.

Then, to. in a recent study, Peck (1977) claims that there exist certain modes of conversation in which the child learner participates, but the adult learner is unlikely to experience. Specifically, Peck considers language play and language during play and suggests that in a play situation the intense emotional climate is one which might foster learning and provide for ample practice which the adult learner is denied.

Thus, in at least three areas, researchers using a discourse analysis framework are making interesting claims.

In a more practical vein, researchers such as Daden (1975) have advocated that conversational analysis be exploited in the development of materials and teaching strategies for the classroom. A more specific recommendation comes from Schwartz (1977) who feels that it is necessary to look at the interaction between second language learners to see if they do indeed adhere to native-like rules in conversation thus testing the popular contention that group work is an effective activity for practicing the target language.

Then, too, a number of recent studies (Allwright, 1975; Fanselow, 1974) have called attention to the role of teachers' corrections of learners' errors in second language instruction. Although we cannot yet, or maybe ever, be prescriptive about the optimal strategy of error correction, perhaps studies of the reaction of the target language speakers to the second language learners' errors might help us to better understand the development in the learner of an awareness of the norms of correctness (Chaudron, 1977).

As we learn more about techniques which initiate and keep conversation going, Hatch (1976) suggests that we might teach conversational strategies to learners. Being able to engage native speakers in conversation will enable a learner to gain more exposure to the target language and perhaps foster more positive attitudes towards target language speakers. For instance, we could perhaps instruct the learner to talk about a small number of topics about which they are prepared to speak. Along the lines of discourse strategies, learners could perhaps be taught the elicitation techniques to get a topic clarified. And finally, Hatch advises, learners should be taught not to give up—that most native speakers will usually try very hard to understand and keep things going themselves if the non-native speaker appears to be really trying to engage them in conversation.

While I certainly have not covered all the areas of second language acquisition being treated through discourse analysis, hopefully I have accomplished what I intended—to offer discourse analysis as an interesting and fruitful approach to the study of second language acquisition from which both research and practical benefits can be realized.

REFERENCES

Allwright, R. L. 1975. Problems in the study of the language teacher's treatment of learner error. In M. Burt and H. Dulay (eds.), *New directions in second language learning, teaching and bilingual education*, Washington, D.C., TESOL.

Arthur, B. 1977. Native speaker input to adult foreign learners of English. Paper presented at 1977 TESOL Discourse Analysis Colloquium.

Celce-Murcia, M. 1977. English language pragmatics and ESL. Paper presented at 1977 TESOL Discourse Analysis Colloquium.

Chaudron, C. 1977. A descriptive model of discourse in the corrective treatment of learners' errors. *Language Learning 27*, 1, 29–46.

Cosaro, W. Undated manuscript. Sociolinguistic patterns in adult-child interaction. Bloomington, Indiana University.

Daden, I. 1975. Conversational analysis and its relevance to teaching English as a second language. Unpublished M. A. thesis, UCLA.

Dulay, H., and M. Bart. 1973. Should we teach children syntax? *Language Learning*, 23, 2, 235-52.

Fanselow, J. 1974. The treatment of error in oral work. Paper presented at 1974 TESOL Convention.

Fillmore, L. 1976. The second time around: cognitive and social structures in second language acquisition. Ph.D. dissertation, University of California, Berkeley.

Gaskill, W. 1977. Correction in adult native speaker—non-native speaker conversation. Unpublished M.A. thesis, UCLA

Godfrey, D. 1977. Oral composition of non-native Adults. Paper presented at 1977 TESOL Discourse Analysis Colloquium.

Hakuta, K. 1975. Becoming bilingual at age five: the story of Uguisu. Unpublished thesis, Departmet of Psychology and Social Relations, Harvard College.

Halliday and Hassan. 1976. *Cohesion in English*. London, Longman.

Hatch, E. 1976. Discourse analysis and second language acquisition. In Evelyn Hatch (ed.), *Second language acquisition*, Rowley, Massachusetts, Newbury House. In press.

Hatch, E., and J. Wagner-Gough. 1974. Second language acquisition. In M. Celce-Muria (ed.), *An introduction to the teaching of english as a second language*. Prepublication manuscript, UCLA.

Keller-Cohen, D. 1977. Systematicity and variation in the development of non-native conversation. A paper presented at the 1977 TESOL Discourse Analysis Colloquium.

Larsen-Freeman, D. 1976. An explanation for the morpheme acquisition order of second language learners. *Language Learning 26*, 1, 125–134.

Peck, Sabrina, 1977. Play in child second language acquisition. Unpublished M.A. thesis, UCLA.

Richards, J. 1971. Error analysis and second language strategies. *Language Sciences, 17*.

Sacks, H., M. Schegloff, and G. Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language 50*, 696–735.

Schegloff, M., G. Jefferson, and H. Sacks. 1977. The preference for self-correction in the organization of repair in conversation. *Language 53*, 2, 361–382.

Schwartz, J. 1977. eRpair in conversations between adult second language learners of English. Unpublished M.A. thesis, UCLA.

Scollon, R. 1974. One child's language from one to two: the origins of construction. Ph.D. dissertation, University of Hawaii.

Selinker, L. 1972. Interlanguage. *International Review of Applied Linguistics 10*, 3.

Vanderbrook, S., K. Schlue, and C. Campbell. 1977. Discourse and second language acquisition of yes/no questions. Paper presented at 1977 TESOL Discourse Analysis Colloquium.

Wagner-Gough, J. 1975. Comparative studies in second language learning. CAL-ERIC CLL Series on Languages and Linguistics, No. 26. Center for Applied Linguistics.

Discourse and Second Language Acquisition of yes/no Questions

Sue Vander Brook, Karen Schlue and Cherry Campbell

University of California, Los Angeles

In this study we examine why second language researchers have trouble determining when learners acquire inverted and statement forms of yes/no questions. We show that the forms of affirmative yes/no questions (inverted or statement with varied stress patterns and optional ellipsis) and their functions (understanding confirmation, encouragement, information elicitation, topic change, etc.) are dependant upon whether the speaker has a greater or lesser degree of presupposition of a *yes*-answer. An analysis of the yes/no questions and accompanying discourse uttered by Spanish-speaking acquirers of English in hour-long, monthly conversations with native speakers over a year show that it is impossible to know when the non-native speaker has internalized the interrelationship of form, function and presupposition. One cannot know if his forms of yes/no questions correctly reflect the amount of presupposition he has and therefore, one cannot determine whether acquisition has actually taken place.

The purpose of this study is to examine a problem in second language acquisition research: why is it that second language researchers have trouble determining when learners acquire inverted and statement forms of yes/no questions? We use the terms inverted and statement form when talking about the surface form of a yes/no question. By inverted form, we mean those questions with the auxiliary element (do-support, modals, auxiliary *have* or *be*) preceding the subject NP, such as 'Do you like artichokes?'. By statement form, we mean those which have the surface form of a declarative sentence but with a rising intonation, for example, 'You like artichokes?'. For the sake of simplicity, this study does not deal with negative yes/no questions or tag questions. (See Bolinger, 1957 for a description of the range of intonation and yes/no questions in English.)

The acquisition of a number of different morphemes has been studied in second language acquisition research in recent years, and it had been assumed that the acquisition of inversion in yes/no questions would follow similar lines as the acquisition of other morphemes. As part of a project carried out at Harvard by Cazden, Cancino, Rosansky and Schumann (1974), the researchers examined non-native speakers' acquisition of English interrogatives. It was hypothesized that in both yes/no questions and WH-questions a progression from uninverted to inverted forms would be observed. A developmental progression of this sort was found for WH-questions, but no developmental pattern could be discerned in the acquisition of yes/no questions. The researchers had operated under the assumption that inversion in yes/no questions was largely optional but that transposition of subject NP and auxiliary would probably increase during the course of acquisition without ever reaching 100%. Their findings, however, did not support this hypothesis. In fact, the percentage of inversion fluctuated up and down from session to session; no increase in inversion was evident over time. The researchers concluded that in English, inversion in yes/no questions is not optional, but that it is probably dependent on contextual and situational variables. It was felt that if the discourse conditions which required inverted and statement forms could be defined, then perhaps a developmental pattern for the acquisition of yes/no questions could be discovered.

Salama (1973) specified seven environments in which statement form questions can be used, but concluded that the production of this form is both a matter of idiolect, in that some people use it more than others, and also a matter of what she calls "mental set", in that an individual may use it at one time very frequently and at another time very infrequently.

Schlue (1975) examined speech data from taped television shows and from the transcripts of the White House Tapes. On the basis of this data, she developed hypotheses about when the statement form is appropriate and when it is not. She then tested these hypotheses on native speakers of English. She found that three of Salama's seven categories were sufficient to classify all the statement form questions found in her data.

Independently from Schlue, as a prelude to analyzing the acquisition of yes/no questions in English by a second language learner, Vander Brook (1975) defined some discourse conditions which control inverted and statement form yes/no questions. She pointed out that, while Salama's work defined several statement form question environments, it did not indicate that the environment can make either the inverted or statement form obligatory.

Yes/No Questions in Discourse

When this study began, the intent was to formalize the constraints on inversion in such a way that we could return to the Cazden et al. study and see how inverted and statement forms are acquired. We began by arguing that if language is rule-governed behavior, then the choice between inverted and statement form should be systematic and therefore predictable. Thus we defined conditions under which one would expect inverted form and conditions under which one would expect statement form, and were able to account for a certain amount of the data. However, the rest did not follow our predictions. If the hypothesis that all aspects of language are rule-governed is correct, we had failed to find the right generalizations to account for all the data. In this case, an examination of a larger corpus of natural language data should reveal the systematicity. On the other hand, the hypothesis itself might be incorrect. In that case, the choice between inverted and statement form is non-systematic and a true optionality of form exists. The fact that we perdicted the form in some cases would be due to chance under this view.

A third alternative, which we support, is that it is not meaningful to regard this area of language strictly in terms of systematicity. Our work shows that the main factor involved in determining the form of yes/no questions is not binary, but is best seen as a continuum. We would make the general claim that when there are non-binary factors controlling a process in language, only when the factors are at extreme ends of their continuums is it meaningful to speak of the behavior in question as systematic or predictable. Because when they are at extreme ends, it is as if they *were* binary. When they are not at the extremes, a particular behavior does not necessarily reflect a quantifiable amount of the factors. We hope to illustrate this position in our discussion of yes/no questions form below.

Interplay of foreknowledge and relevance. It seems that what a speaker says relates to the previous speaker's utterance and is relevant to it in both speakers' minds. Thus a speaker doesn't say just anything, or even anything that may be related in his mind to what the previous speaker has said, but makes sure his utterance will be relevant from the previous speaker's view point too. This kind of pragmatic judgment of appropriateness is reflected in the form an utterance takes.

Assume two speakers, A and B, have just met at a party. They share a great deal of knowledge and beliefs about the world and each other. But there are many details that they don't share. To see how this affects the form of their utterances, let us assume the following exchange takes place:

(1) A: My husband and I are going on vacation next week.

B: 1. Where are you going?
2. Are you going down South?
3. You're going down South?
4. You're going down South.

First note that any of the possible utterances B1-4, by repeating a part of A's statement (subject + verb) would appear to relate to it. But repetition alone need not be equivalent to relevance. In choosing the most relevant response from B1-4, B's choice will be affected by his knowledge about A. Thus if we were observers of this conversation and heard A's statement followed by B1, we would assume B didn't know the answer to his own question. Moreover, we would assume that A also knows that B's knowledge about him does not include the answer to that question. If this is the case, B1 is a relevant response in both speaker's minds.

Alternatively, if we hear B follow A's statement by B2, we would be led to believe that *down South* is a relevant place for A to go and that A will think that *down South* is a relevant place to go. B2 would thus be relevant in spite of the fact that they had just met if certain facts were shared knowledge. This might be the case if this conversation took place in the midst of winter in a frigid Northern city and the news is full of how people are fleeing to the South to escape the rigours of the cold.

If A's statement were followed by B3, or B4 we would believe that B has an even greater knowledge about A and that both parties know it. This would occur when, for example, B and A have had previous contact and they both know that A wants to visit the South when the first opportunity comes along.

In the above exposition we do not mean to imply that only one form (B1, 2, 3, or 4) is necessarily the only relevant utterance in any given situation. In some cases, more than one form may be appropriate. For example, where we suggested B2 could be a relevant utterance to follow A's statement, B1 might also have been used. However, the use of B1 in a situation where B4 would be appropriate could only be done in a humorous vein with B pretending ignorance.

We have noted a succession and overlap of one form to the next as they related to shared knowledge. It is interesting to note that there is a parallel overlap of certain formal properties of these four utterances as well: B1 and B2 are similar in that they are both inverted, B2 and B3 in that they both have rising intonation, and B3 and B4 in that they are both statement form.

The remainder of this paper will be concerned with the domain of the yes/no question, that is, those situations where a WH-question or a statement would not be relevant. When affirmative yes/no questions are posed, the speaker has varying degrees of certainty that the interlocutor will answer affirmatively. We can refer to this presupposition of a *yes*-answer as a continuum ranging from low to high.

High presupposition of a *yes*-answer. Now let us consider some possible conversations between A and B, in which B holds high presupposition of a *yes*-answer. We will examine the forms of the question that reflect this presupposition, and the functions they fulfill in the discourse.

The first is one in which B's utterance essentially echo's A's, the so-called "echo" question.

(2) A: I'm studying poetry this term.
B: You're studying poetry this term?

We know B must have high presupposition of a *yes*-answer because of what A has just told him. Thus while B is not asking for information at all, his utterance could have other functions. The utterance could be a means for B to confirm his understanding of A's statement. It could indicate that A's statement contains new information or that B's previous knowledge was contradicted by A's statement. It could display B's interest in the information in A's statement. It could allow or encourage A to continue. (This list of functions may not be exhaustive and certainly does not apply exclusively to this form of utterance as a consideration of WH-questions, embedded questions or statements will show.)

Now consider a correspondingly neutral-stressed inverted-form question in the same situation as above.

(3) A: I'm studying poetry this term.
B: #Are you studying poetry this term?

(The symbol, #, indicates the sentence following it is a grammatical one but is inappropriate in the particular discourse in which it is shown). This inverted form question fails to acknowledge A's statement. It signals low presupposition of a *yes*-answer when one would expect B to have high presupposition; thus it is not a relevant response from A's point of view. The intonation contour on a WH-question also affects the relevance of the particular question to a given discourse:

(i) A: I'm studying something unusual this term.

B: 1. What?
2. What.

If we heard A's utterance followed by B1, we would assume B knew A had spoken but had not understood all or part of what A said. On the other hand, B2, without falling intonation, would indicate that B understood A's statement. It would not however be an appropriate response if there were no unspecified NP in A's statement as is shown below:

(ii) A: I'm studying poetry this term.

B: 1. What?
2. #What.

The inverted form is acceptable if the stress is shifted:

(4) A: I'm studying poetry this term.
B: Aré you (studying poetry this term)?

(The diacritic, ′, indicates the position of nuclear stress, while, ″, indicates heavy or emphatic stress). Here, the auxiliary element carries the main stress. As observers of this exchange, we know that B has understood A. Furthermore, the same range of functions possible on the statement form questions in (2), are equally possible in this stress-shifted inverted form.

The same stress-shift to auxiliary element is possible on the statement form:

(5) A: I'm studying poetry this term.
B: You are (studying poetry this term)?

As in (4) and (2) above, we know that B has understood A, and the same range of functions is still possible.

In the statement form, the stress could also have shifted to the subject NP:

(6) A: I'm studying poetry this term.
B: Yóu're studying poetry this term?

or Yóu (are studying poetry this term)?
or Yóu are (studying poetry this term)?

Again this stress-shifted form indicates that B has understood A, and it may carry the same range of functions.

In the inverted form, shifting the stress to the subject NP seems to be restricted to contrastive usage as in (i) and (ii):

(i) A: I know Jane's studying poetry this term, but are yóu (studying poetry this term)?
(ii) A: Jane's studying poetry this term.
B: And are yóu (studying poetry this term)?

In neither case does B have a high presupposition of a *yes*-answer. Because we haven't dealt thoroughly with this contrastive use of stress, we haven't included it in Figure 1.

In addition to the functions discussed with respect to example (2) above, there is one other that deserves special consideration. Any one of the questions in examples (4), (5) and (6) can be used to express B's opinions, emotions and assessments towards A's statements, including doubt, surprise, skepticism, challenge, disbelief, outrage or reversal of contradictory presupposition. The strength of any one of these may vary. These factors are grouped together because they are all expressed by varying amounts of phonological changes including syllable lengthening, exaggerated intonation rise and increased volume. We use the word EXPRESSION as a cover term for these factors. Thus the inverted question in (4) with stress shifted to the auxiliary, when conveying EXPRESSION, might be read:

(7) AR̋E:: you (studying poetry this term)?

The facts covered in this section are graphically displayed on the right-most portion of Figure 1 which shows the acceptable forms of yes/no questions when presupposition of a *yes*-answer is high.

Low presupposition of a *yes*-answer. The form of yes/no questions used to communicate low presupposition of a *yes*-answer is different from that used to express high presupposition of a *yes*-answer. Moreover, the range of meanings is restricted when presupposition is less.

First, the actual words used in B's questions are not the same ones as in the previous speaker's utterance. None of the statement forms are permissible, nor are any of the stress-shifted forms (but see discussion of example (6) above), and as a result EXPRESSION cannot be conveyed. The only form we have found is a neutral inverted form. Situations in which the questioner has little presupposition of a *yes*-answer include those where he does not know the answer and is eliciting information, ('Did you take out the garbage?'), polite commands or requests, ('Would you mind shutting the door?', 'Do you know what time it is?'), topic change or attempts to get the floor, ('Did you remember the girl who used to live upstairs?'), etc. Consider the following example:

FORMS:	A: . . . job opportunities . . . B: Here we are; thanks for the ride. Uh . . .	A: We just got back from vacation.	A: I'm studying poetry.
neutral inverted (optional ellipsis of AUX)	Would you like to come in for a drink?	B: Did you have a nice time?	
stress-shifted inverted (to AUX)			B: Are you (studying poetry)?
neutral statement		B: You had a nice time?	B: You're studying poetry?
stress-shifted statement (to subject NP)			B: You're studying poetry? B: You (are studying poetry)? B: You are (studying poetry)?
stress-shifted statement (to AUX)			B: You are (studying poetry)?

LOW ← PRESUPPOSITION OF A YES-ANSWER → HIGH

FUNCTIONS:

information elicitation →
topic change
clarification of previous inappropriately answered yes/no question
asking a directed question

← understanding confirmation
acknowledgement of new or contradictory information
display of interest
encouragement
EXPRESSION

Figure 2. Interaction of Forms and Functions of Affirmative Yes/No Questions.

(8) A: . . . job opportunities . . .
B: Here we are; thanks for the ride. Uh . . .
Would you like to come in for a drink?

In example (8) we shall assume A is dropping B off at his home and on the way they've been talking about job opportunities. As they arrive at B's home, he asks if A would like a drink. In this situation B has low presupposition of a *yes*-answer, and can use the neutral inverted form. It changes the topic and/or asks for information. The same question would be inappropriate if asked in any statement form or stress-shifted inverted form.

(9) A: . . . job opportunities . . .
B: Here we are; thanks for the ride. uh . . .
#You'd like to come in for a drink? (neutral statement)
#Yóu'd like to come in for a drink? (stress shifted statement)
#You wóuld like to come in for a drink? (stress shifted statement)
#Wóuld you like to come in for a drink? (stress shifted inverted)
#Would yóu like to come in for a drink? (stress shifted inverted)

Each of these forms is inappropriate because it makes an incorrect assumption about what is shared knowledge.

Along with information elicitation and topic change, another function requiring the neutral inverted form is clarification of a previous inappropriately answered yes/no question. If A and B are conversing under conditions such that A replies to B's question inappropriately, (such as over a poorly connected telephone or at a loud party), if B continues to ask the question at all, he will likely choose a neutral inverted form. He abandons any presuppositions carried in his earlier question.

(10) A: We just got back from vacation.
B: You had a nice time?
A: We got back at 4:30. (inappropriate response)
B: Oh, but did you have a nice time? (clarification, neutral inverted)
A: Yeah, it was nice.

Number (11) below is our last example of a neutral inverted question and in this case it is when a person is instructed to ask another person a yes/no question. The person who asks the directed question acts as though the topic were new and as if he had low presupposition of a *yes*-answer. This situation is created in a second language classroom or a speech data gathering situation in which the object is to elicit yes/no questions.

(11) Teacher: Ask Roberto if he saw *Cannon* last night.
Student: Did you see *Cannon* last night?
Roberto: Yes.

It would be inappropriate to respond to the teacher's directive with a statement form ('You saw Cannon last night?') even if the student and everyone else in the class knew what Roberto did the previous evening, because the whole activity is understood to be role-playing, in this case for language learning.

The information shown in examples (8-11) and accompanying discussion are summarized on the left-most portion of Figure 2. (Again, the list of functions may not be exhaustive and does not apply exclusively to this form of utterance.) Both in function and form, the appropriate question in the low range of the continuum differs from those in the high range. In particular, the words B uses with low presupposition of a *yes*-answer are different from those of A that they follow, while those of B with high presupposition are essentially identical to A's. Furthermore, the neutral inverted form required in situations of low presupposition, is not possible in those of high presupposition, and likewise, the statement with stress-shifted forms required in situations of high presupposition are not possible in those of low presupposition.

Intermediate Presupposition of a *Yes*-Answer

Now let us consider that part of the continuum where presupposition of a *yes*-answer is somewhere between the low and high ranges. These are cases in which B is rather certain that A will answer affirmatively, not because of what A has just said as in the high range, but because of B's beliefs about reality that he and A share. There are two ways B can phrase this question and they are illustrated in (12):

(12) A: I just got back from vacation.
B: 1. Did you have a nice time?
or 2. You had a nice time?

Note that B's questions are not a syntactic echo of A's statement and in that regard they resemble the utterances in the low range.[1] The neutral inverted form reflects one function of the low range, that of eliciting information. The use of the neutral statement form reflects one of the functions of the high range, that of confirming understanding. This is indicated by the arrows next to each of these functions, extending them further along the continuum. While in the cases discussed under the high range, B was confirming what A had just said, in example (12) B is checking an unstated assumption. We believe that B's use of this form indicates that he has a strong presupposition of a *yes*-answer, perhaps from his previous knowledge about A, from his experiences with vacations in general, or from his accumulated beliefs about reality. In a sense, the use of this form equates "nice time" to "vacation" in A's statement. However, we cannot predict from our own intuitions that B has this equivalency in *his* mind.

---

[1] Note further the problem that this poses for transformational grammars (see Stockwell, Schachter and Partee, 1973) which would derive (12)2 'You had a nice time?' from an underlying form 'Did you say you had a nice time?' by an optional transformation, where 'Did you say you had a nice time?' would be an inappropriate response to A's utterance 'I just got back from vacation'.

Furthermore, if he used the neutral inverted form, we can only presume that he does not equate "nice time" with "vacation". We cannot be certain. Possibly some other factor keeps him from using the statement form—not wanting to sound presumptuous, for example, or wanting to give a false impression of his presupposition. Therefore, in this intermediate range of the continuum there need not be a one-to-one correspondence of form to presupposition. Furthermore, in this range, it is no longer meaningful to regard the choice of form as systematic or predictable. Giving a false impression of presupposition would constitute a manipulative use of language. Thus a speaker can mispresent reality by using one form of a yes/no question where the other would have been more truthful. We can illustrate this sort of manipulative use of language by referring to example (1) which involved two speakers who had just met. For convenience the relevant parts are repeated here:

(1) A: My husband and I are going on vacation next week.
B: 2. Are you going down South?
3. You're going down South?

We pointed out that the use of B2 would be appropriate if certain facts concerning the hard winter weather were shared knowledge. We said further that B3 would be appropriate if A and B had had previous contact and knew each other well enough for B to be more certain that A would be going South given the opportunity. Now notice how B's choice of form can mislead a co-participant C, who does not know A, into thinking that B knows A better than he does. B may think that it is highly likely that A will go South because of the bad weather and because he wishes to impress C, uses B3, the statement form question, thereby sounding as if he has "insider's knowledge". If A responds affirmatively, then C may fall for B's little ploy. (But if A replies that they are not going South, then B's gamble may or may not succeed. Sue Gass pointed out in personal communication that C may still believe B had more knowledge of A than he does simply by virtue of the fact that B has used a more highly presuppositional form). On the other hand, if B knows A very well, but wishes C to think he doesn't, he may use the inverted form which implies less presupposition of a *yes*-answer, and he may lead C to think he doesn't know A well.

The statement form is so powerful a conveyer of presupposition that it can seemingly be used to manipulate the other speaker into giving the desired answer, even when that answer is contrary to fact. In a series of Doonesbury cartoons, Ms. Caucus was crushed to find that Andy would not return her romantic feelings. Andy however, wanted to maintain the friendship. In the face of knowing that she felt uncomfortable with him, he asked her, "We're still cool, then?" Ms. Caucus felt forced to respond, "Of course." Then she added, "Go away." indicating her true feelings (Gary Trudeau, March 10, 1976). (This type of manipulation is what conversational analysts (Schegloff, in 1977 UCLA lecture) describe as structuring a first pair part to receive a specific preferred second pair part.)

Returning to the distribution of question forms with intermediate presupposition of a *yes*-answer, it should be noted that none of the stress-shifted forms would be possible in example (12):

(13) A: I just got back from vacation.
B: #Yóu had a nice time?
#You hád a nice time?
#Did yóu have a nice time?
#Díd you have a nice time?

Stress-shifted forms imply that the neutral form is already part of the discourse which is the case in the right-most portion of Figure 2 where B's utterance repeats A's words, but is not the case in the example (13). In spite of a possible semantic similarity between A's statement and B's question in example (13) the words of B's question are not yet in fact part of the discourse until he utters them. The fact that the stress-shifted forms are not possible, means, moreover, that EXPRESSION, is not possible.

This discussion is summarized and appears in the center-most portion of Figure 1. The whole of Figure 1, then, is a graphic representation of the interrelationship between presupposition of a *yes*-answer, form, and function of a question.

Schlue (1975) tested the claim that statement form yes/no questions are used to question an assumption or verify a hunch. She found that stress shift combined with the phonological changes mentioned above which denote EXPRESSION, are very powerful conveyors of high presupposition of a *yes*-answer. Furthermore, the neutral statement form also conveys an expectation of a positive response. Native speakers were asked to listen to a tape containing contextualized yes/no questions of this sort:

(14) a. Are you planning to hire another waitress? (neutral inverted)
b. You're planning to hire another waitress? (neutral statement)
c. You're planning to hire anóther waitress? (stress-shifted statement)

The subjects were then asked to interpret each question according to the message thy heard:

(A) The speaker has little or no idea whether the answer will be *yes* or *no*.
(B) The speaker has a hunch that the answer will be *yes*, and is testing that hunch.
(C) The speaker has a hunch that the answer will be *yes*, and is registering surprise, disbelief, or outrage.

The type represented by (14a) (neutral inverted) received the expected response (A) 80% of the time; type (14b) (neutral statement) received the expected response (B) 74% of the time; and type (14c) (stress-shifted statement) received the expected response (C) 97% of the time. (Schlue speculated and offers some evidence that departures from the expected interpretations could be attributed to various combinations of faulty item construction and individual speaker differences.)

Ellipsis of *Have*, *Be* and Do-Support

In spoken English, the auxiliaries *have* and *be* and do-support often do not appear in yes/no questions.[2] This is illustrated in the following examples involving second person subjects.

(15) You ever been to San Francisco?
(16) You gonna travel during vacation?
(17) You wanna go to the store?
(18) You lose something?

When observing the usage of the various forms of yes/no questions in speech data, the problem arises whether such questions should be considered as being elliptical questions in inverted form or questions in full statement form. A question such as (17) above gives no indication of the related form.

(19) a. Do you wanna go to the store?
b. You wanna go to the store?
c. You wanna go to the store.

It could be that the auxiliary element, *do*, has been elided from its initial position in the equivalent inverted question (19a). Or it could just as well be that the *do* never appeared at all because (19b) is in statement form and is related to a statement in which do-support does not surface (19c). Consider, however, the following examples in which there is syntactic evidence to support a claim that such questions should be considered elliptical inverted.

(20) a. Have you ever been to San Francisco?
b. You ever been to San Francisco?
c. *You ever been to San Francisco.
(21) a. Are you gonna travel during vacation?
b. You gonna travel during vacation?
c. *You gonna travel during vacation.
(22) a. Did you lose something?
b. You lose something?
c. *You lose something.

If yes/no questions in statement form are related to grammatical statements, then the questions (20b, 21b, and 22b) above cannot possibly be related to statement form questions since the related statements (20c, 21c, and 22c) are ungrammatical. We would argue that the derivation of all such questions should be the same, relating them to the corresponding inverted form and not to the statement form yes/no questions.

---

[2] Schmerling (1973, p. 580) also notes that *have*, *be* and do-support can be elided in yes/no questions. She discusses ellipsis of the subject NP in such questions as well. We acknowledge its occurrence but will only discuss ellipsis of the auxiliary element insofar as it relates to the forms and functions of yes/no questions presented above.

There is further evidence to support the claim that such questions are elliptical-inverted. Consider the distribution of the (b) questions in discourse; it is the same as that of the (a) questions. (Schlue, 1975 notes this and it has repeatedly been substantiated in the work being reported on here.) Consider a situation in which a wife walks into the living room on a Sunday afternoon and addresses her husband.

(23) a. Are you tired of sitting around?
b. You tired of sitting around?
c. #You're tired of sitting around?

Both questions (a and b) are appropriate to this situation. The statement form question (23)c is not used to change topics (a function represented in the low range of presupposition), whereas both the inverted question in full form (23)a and the question with the elided auxiliary element (23)b do so. This semantic, or discourse-related justification, coupled with the syntactic justifications, discussed above, is a strong indication of the equivalence of these questions to inverted questions in full form. They are then elliptical yes/no questions, with the inverted auxiliary element elided.

It should be noted that there are no elliptical counterparts for the stress-shifted inverted form question in the high range. When the nuclear stress falls either on the auxiliary or on the subject NP of a yes/no question, the auxiliary element cannot be elided. Ellipsis of the auxiliary element cannot occur simply because a constituent in a sentence that carries the stress cannot be elided. Thus the inverted question in the high range (here example 24) cannot take on a form as in (25) without any nuclear stress at all:

(24) Are you (studying poetry this term)?
(25) * You (studying poetry this term)?

The following inverted yes/no question is one in which the nuclear-stress is contrastive and falls on the subject NP.

(26) I know Jane wants to go to the store, but do yóu want to (go to the store)?

If the *do* were elided the resulting form would be the same as the statement form of the question. The statement form could equally well be used, but it would represent a different set of functions (those of high presupposition) than the inverted question above (those of low presupposition). In sum, only the *neutral* inverted form questions seen in the low range of the continuum permit ellipsis of auxiliary elements.

Finally, it is still unclear as to whether all forms of *have*, *be* and do-support can be elided in all combinations of variables such as stative/non-stative verb, present/past context, first/second/third person, or pronominal/full subject.[3] Campbell is currently testing the ellipsis of do-support in yes/no questions to see which of these factors affect its acceptability among native and non-native speakers.

Methodology[4]

The data used in this study were collected by Cazden, Cancino, Rosansky and Schumann (1975) and are that of two of their six subjects. The subjects were native speakers of Spanish. They were largely free learners acquiring English by exposure to the English-speaking environment and without formal instruction. They were visited approximately once every two weeks for a ten-month period. During these hour-long visits the researchers gathered both spontaneous and experimentally-elicited speech.

We began our investigation of the data by independently pulling the yes/no questions from the even-numbered transcripts. We then evaluated the corpus independently in the following manner. First we attempted to determine the question forms a native speaker would use in each discourse environment. Next we compared our so-called appropriate forms with the forms actually used by the non-native speaker. Then we came together and compared our judgments and tried to settle on one judgment for each item. Finally, we listened to the tapes in order to clarify our opinions in different cases.

Second Language Learner Data

Initially during our analysis, some questions and their environments in the data seemed clear-cut and the evaluation procedure was rather straight-forward. Consider the following:

(27) (E = native speaker, C = non-native speaker)
E: . . . Okay, good. Wanna go see the mummies? (in a museum)
C: What (are) mummi(ies)?
E: Do you know what a mummy is?
C: No.
E: Well, it's a guy who died a long time ago. And they kept him all wrapped up and you can still see him. Sound interesting?
C: xxx xxx xxx xxx
E: Right.
→ C: Hey, is this a house?
E: It's a desk. See you put all the papers in here and notes and letters.

This example seems to display the type of question used to reflect low presupposition of a *yes*-answer. There is previous talk about mummies (equivalent to the previous talk about job opportunities in the low range) and then apparently the learner (C) wants to ask a question that changes the topic and/or ellicits information. To accomplish this a native speaker would use the neutral inverted form, as did the learner.

---

[3] Schmerling (1973, p. 580) states that ellipsis of the auxiliary element is not possible with the past tense of *have* and *be* but that it is possible across all three persons. We, however, have not examined enough data to be able to support these claims.

[4] John Schumann joined the authors in the investigation and analysis of this data. His help and guidance are gratefully acknowledged.

The next sequence seems to exemplify the other end of the continuum, that is, high presupposition of a *yes*-answer.

(28) E: Did you see that building in Boston? The one that's all glass like that? I live near there.
→ C: You live there?
E: Yeah, near there. Near, around, in the back.

The learner's question is a near repetition of the experimenter's (E's) preceding remark. The learner is likely checking his understanding of the previous utterance. In this case, a native speaker could use a statement form, as did the learner.

Difficulties in our analysis procedure first arose with examples such as the following:

(29) (discussing camels)
E: And doesn't have, it, drink, it can go two weeks with no water. See, it drinks alot of water and puts it all up here in that hump on the top, and it saves water.
→ C: Yeah. And it, it, it can eat some gasoline, too?
E: It doesn't eat . . . It'd die from the gasoline.

Here the learner's yes/no question is in no way a repetition of a previous utterance by the experimenter. It is therefore not strictly a part of the high range. But we cannot define the amount of presupposition any further. We cannot know how strong the presupposition is that the experimenter will agree that camels can store gasoline as well as water in their backs. His presupposition may be weak; he may truly be asking whether this information is true or not. In this environment a native speaker would use a neutral inverted form, thus the learner's actual statement form would be inappropriate. The learner's presupposition, on the other hand, may be strong; it could well be a logical conclusion drawn from the information the learner has just heard about the capabilities of camels, coupled with his prior knowledge about automobiles. He would then have reason to assume that the experimenter will affirm the proposition and he would be checking that unstated assumption. In this case, a native speaker would use the neutral statement form as did the learner. In deciding which question form is appropriate to the discourse environment, it is necessary to know how much presupposition the learner has about the answer and it is impossible to measure that amount of presupposition from the discourse environment. Without knowing what presupposition is represented, it becomes impossible to determine whether the form of the question that the subject produces is appropriate or not.

Not only can we not judge the appropriateness of the question in (29), (and exchanges like this comprise a large part of the corpus) we now believe one cannot conclusively judge the appropriateness of the learner's form in (27) and (28) either. It was stated above that in example (27) the learner likely has low presupposition. The main reason for drawing this conclusion is that the question used by the learner here is in neutral inverted form, which reflects low presupposition. However, it could just as well be that the learner sees the object and really does believe that what he sees is some kind of a house. His presupposition of a *yes*-answer might indeed be very high and the neutral inverted form that he used would fail to reflect his true amount of presupposition. In this case he would have made an error and formed his question inappropriately. Similarly, the learner might have had some reason in example (28) to have low presupposition, e.g., perhaps he did not hear or understand when the experimentor said that he lived there. Maybe the learner really did want to ask a question eliciting information. If this were the case, he used the wrong form.

The requirement of *relevance* in discourse permits one to use the form to get at the amount of presupposition of a *yes*-answer in a native speaker of English. It was on this basis that Figure 2, depicting the interrelationship of form and presupposition, was developed. With non-native speakers, however, there is no variable that reflects whether the non-native speaker has internalized the interrelationship of form and presupposition. One cannot know if his forms of yes/no questions correctly reflect the amount of presupposition he has. Thus, without a means of measuring presupposition independently from form, we cannot say which form a non-native speaker should use. For these reasons then, second language researchers cannot determine when the acquisition of inverted and statement form occurs in second language learners.

REFERENCES

Bolinger, D. L. 1957. *Interrogative structures of American English (The direct question).* Publication of the American Dialect Society, No. 28. University, University of Alabama Press.

Cazden, C., H. Cancino, E. Rosansky, and J. Schumann. 1975. *Second language acquisition sequences in children, adolescents and adults.* Final report, National Institute of Education (Grant #NE-6-00-3-0014).

Salama, N. 1973. An analysis of the occurrence of the statement question form in spoken English. *English Language Teaching* 27, 3, 262–265.

Schlue, K. 1975. Yes/no questions in English: inversion vs. statement form. Unpublished paper, TESL department, UCLA.

Schmerling, S. F. 1973. Subjectless sentences and the notion of surface structure. *Chicago Linguistic Society Papers from the 9th Regional Meeting,* April 13–15, 577–586.

Stockwell, R. P., P. Schachter, and B. H. Partee. 1973. *The major syntactic structures of English.* Holt, Rinehart and Winston, Inc.

Vander Brook, S. 1975. How about this one? Acquisition of yes/no questions by a second language learner. Unpublished paper, Linguistics Department, UCLA.

Conscious Communication Strategies in Interlanguage: A Progress Report

Elaine E. Tarone

University of Washington

Recent studies in the field of second language acquisition on the nature on "interlanguage" (Selinker, 1972) have made reference to "strategies" which are hypothesized to be used by the learner in the attempt to both understand and produce utterances in the target language. Of particular interest in this study is the nature of the communication strategies of production which are used by the learner in situations where he or she is unfamiliar with, or uncertain of, the appropriate target language form or rule. This study represents an attempt to examine such communication strategies of production in more detail, using the terminological framework developed in Tarone, Cohen & Dumas (1976).

The speech production of adults from a variety of first-language backgrounds, learning English as a second language at a large university, is examined as they perform a picture description task in both their first language and the second language, English. The communication strategies evidenced in their interlanguage are described, using the above-mentioned terminological framework; in addition, an appeal is made to the intuitions of the learners themselves, who are asked to explain why they think they used the strategies they did. The analysis of the data thus obtained is used to determine the extent to which various communication strategies can be clearly differentiated, the extent to which learners seem to rely on preferred strategies, and the correlation of such preferred strategies with the first-language background of the learner, or other learner variables.

The field of second-language acquisition, if it chooses to address itself to the issue, may be able to provide valuable insights into the nature of the interrelationship of language, cognition and personality (cf. Guiora *et al.* 1975). The process of second-language acquisition can, I believe, provide some fascinating insights into the workings of the mind in the struggle to communicate meaning through language—particularly when we are able to study this struggle at the level of the individual. In particular, the strategies used for the communication of a desired concept when the requisite target language (TL) term is lacking, may provide us with important information about the cognitive organization of linguistic information.

I report here on a study which is in progress, and which focuses upon such communication strategies. I would like to outline the parameters of the study, report on the method of data analysis I am developing, and describe some preliminary patterns which are emerging from the data.

Tamas Váradi (1973) has done a study on the nature of what he termed "communication strategies". According to Gorbet (1974), Váradi pointed out the fallacy inherent in looking only at the *form* of a language learner's utterances—the learner herself is aware of her deficiencies in the target language and consciously attempts to compensate for those deficiencies by using communication strategies. Váradi uses the term "communication strategy" to mean a conscious attempt to communicate the learner's thought when the interlanguage (IL) (Selinker, 1972) structures are inadequate to convey that thought. Váradi's central point seems to be that even when the learner produces interlanguage forms which are syntactically correct, she still may not have produced forms which communicate her *intended meaning*.

In essence, then, conscious communication strategies are used by an individual to overcome the crisis which occurs when language structures are inadequate to convey the individual's thought. This crisis may—and indeed does—occur when one speaks in one's first language, as well as when one attempts to communicate in an interlanguage. We have all experienced the situation of trying to "get an idea across" for which we somehow "can't find the words". However, because the interlanguage is a simplified system (Richards, 1975) and also a permeable system (Adjemian, 1976), the use of conscious communication strategies in interlanguage is far more frequent—and in some ways far easier to study, since we have the native language (NL) to draw on as a resource in determining the learner's intended meaning.

Research on the nature of conscious communication strategies (in Váradi's sense) has relied upon studies in which second-language learners are presented with a series of pictures and are asked to describe those pictures in the target language. This technique has succeeded in forcing the learners to communicate about specific content in the interlanguage, while allowing them considerable leeway with regard to the way in which they communicate that content. It also has allowed the investigator to make some assumptions about the speaker's intended meaning. Several studies have been done on the nature of communication strategies. Váradi (1973) isolated three primary types of communication strategy in his work with adults in Eastern Europe learning English as a foreign language. Tarone, Frauenfelder and Selinker (1976) (henceforth TFS) independently did a similar study in Toronto, working with children learning French as a second language, and found a very similar grouping of communication strategies. Hendrickson and others are doing similar work in Ohio.

To my knowledge, in these studies the subjects have only been asked to narrate the stories in the TL—but not in the native language as well. There are obvious advantages to having the same subjects perform the same task in both NL and TL. In order to isolate as clearly as possible the intended meaning—that which the second-language learner is trying to communicate—we need to know what she says when using a language code which provides her with as many language structures—as many options—as possible. What, specifically, is the learner trying to say about each picture frame? Having the learner narrate the story in the NL helps to answer this question. For example, if we have a picture frame where a learner says in NL, the equivalent of "The coffee spilled", but in IL, "The coffee fall down," we have a fairly clear idea that the learner is trying to communicate in IL a more specific content ("spilling") than she has the IL vocabulary for. The study I report on here, then, is one in which the learners perform the same story-telling task in both NL and their English IL.

Subjects and Procedure

The subjects were nine adults learning English as a foreign language at the University of Washington in Seattle; their first-language backgrounds were Spanish Turkish, and Mandarin (with three subjects from each NL background). All nine subjects were at an "intermediate" level of spoken English, although they seemed to very a good deal in proficiency within that classification. (As Larsen-Freeman, 1977, has pointed out, our present inability to find a good index of development for second-language learners makes it very difficult for researchers to make reliable statements about their relative levels of proficiency.) I made a subjective estimate of the learners' relative levels of proficiency in English; the subjects are listed in order of approximate relative proficiency in Table 3.

The subjects were shown two series of simple drawings, and an illustration from a children's book, and then asked to tell the stories depicted by the illustrations, in both their NL and in English. The pictures were chosen to precipitate a crisis of communication in which the subjects were attempting to communicate concepts for which they did not yet know (or were unsure of) the correct target language form. The subjects were recorded as they told the stories. Afterwards, the experimenter reviewed the task with the subject, asking whether the subject knew particular forms in English, or asking why the subject used one form rather than another. The subjects' responses to these follow-up questions were also recorded. Their responses to the picture tasks in both NL and IL were transcribed.[1] Seven of the most difficult target concepts have been selected for the purpose of this report, and the subjects' responses to each image in NL and IL have been tabulated, as in Table 2.

Data Analysis

The analysis of spontaneous speech data is truly frustrating. The language researcher is always tempted to reduce language data to *measurable entities* which can be counted, graphed and statistically weighed, and yet is always aware that this very process can distort the data, to the extent that one must question whether one is still looking at anything remotely resembling *language*. One sets up a system to described human behavior, which fits as accurately as possible what one sees, yet ultimately no system does justice to the reality. In spite of this, we go ahead and set up systems because, despite their limitations, we need them to help us see those patterns which do exist in human behavior.

---

[1] My thanks to Selda Agar and Bob Cornett for their help in translating and transcribing the Turkish and Mandarin data.

Tarone, Cohen and Dumas (henceforth TCD) (1976) attempted to systematize a taxonomy of communication strategies. However, it is clear to me that any such attempt to establish an enlightening typology of clear-cut mutually-exclusive communication strategies is bound to run into trouble as soon as we begin to deal with real data. In wrestling with my data, I have found that I have had to reorganize the TCD and Váradi taxonomies somewhat. The result is a system which seems to provide the best tool to make sense of the behavior of my subjects in this communication situation.

One can, I believe, differentiate five basic conscious communication strategies fairly consistently: avoidance, paraphrase, conscious transfer, appeal for assistance and mime.[2] These strategies are outlined and illustrated in Table 1.

(1) *Avoidance*. There are several types of avoidance strategies available to

TABLE 1

Typology of Conscious Communication Strategies

Communication Strategy	Interlanguage	Target Language
(1) *Avoidance*		
(a) Topic avoidance	ø	mushroom
(b) Message abandonment	the water (mumble)	the water spills
(2) *Paraphrase*		
(a) Approximation	pipe	waterpipe
	labor	work
	lamp	waterpipe
(b) Word coinage	person-worm	caterpillar
	jug-worm	caterpillar
(c) Circumlocution	Something, I don't know what's its name. That's, ah, that's Persian and we use in Turkey, a lot of.	waterpipe
(3) *Conscious transfer*		
(a) Literal translation	He invite other person to drink	they toasted each other
(b) Language switch	balon	balloon
	tirtil	caterpillar
(4) *Appeal for assistance*	What is this? What called?	waterpipe
(5) *Mime*	. . . and everybody say [claps hands]	everybody applauds

---

[2] I am grateful to a participant (unknown to me) in the TESOL 1977 Convention, who uggested this category during a discussion following the reading of this paper.

these learners. *Topic avoidance* (TFS) occurs when the learner simply does not talk about concepts for which the vocabulary is not known. So, if one is unsure about how to say "mushroom" in English, one steers the conversation away from the topic. *Message abandonment* (a term coined by Váradi, and redefined in TCD) occurs when the learner begins to talk about a concept but is unable to continue and begins a new sentence. For example, in Table A, the learner begins to say, "The water spills," but instead her voice trails off and she goes on to the next picture.

(2) *Paraphrase*. A variety of approaches are subsumed under this category. I have altered TCD's definition somewhat, and here define paraphrase as the rewording of the message in an alternate, acceptable target language construction, in situations where the appropriate form or construction is not known or not yet stable.[3] There are three basic types of paraphrase which I have been able to find consistently. (a) The first is *approximation* (discussed in TCD)—the use of a single target language vocabulary item or structure, which the learner knows is not correct, but which shares enough semantic features in common with the desired item to satisfy the learner. Examples of approximation would include the use of "high-coverage words" (Levenston, 1971) such as "worm" for "silkworm", or "pipe" for "waterpipe"; "low-coverage words" (Levenston, 1971) such as "labor" for "work"; or approximations which operate at more or less the same level of generality, but are simply inappropriate, such as "lamp" for "waterpipe". It is clear that approximation may be a type of overgeneralization—it is often only the learner's concious use of the strategy which differentiates it from overgeneralization, as that term has traditionally been used in the literature. That is, the learner *knows* in using approximation that the term is wrong, but uses it anyway; whereas, in overgeneralization, the learner may or may not know the term or rule is wrong. (b) *Word coinage* (Váradi's term) occurs when the learner makes up a new word in order to communicate a desired concept. For example, Váradi reports the use of the term "airball" for "balloon", and my subjects referred to an animated catepillar as a "person worm" or a "jugworm". (c) *Circumlocution* (Váradi's term) is a wordly extended process in which the learner describes the characteristics or elements of the object or action instead of using the appropriate target language structure. (Váradi distinguishes circumlocution and description, but I found it more useful to combine these categories.) For example, in describing a waterpipe, a subject in my study said, "She is, uh, smoking something. I don't know what's its name. That's uh, Persian, and we use in Turkey, a lot of."

(3) *Conscious transfer* may take the form of literal translation or languag switch (TCD). Literal translation occurs in this study when a Mandarin speake describes two persons toasting one another by translating the equivalent Mandarin expression word for word: "He invites him to drink." In language switch, the learner simply uses the NL term without bothering to translate, as, usin

---

[3] TCD maintain that paraphrase occurs when the learner is attempting to *avoid* a term; I no longer believe paraphrase to be a form of avoidance.

"balon" for "balloon." I believe, from the data gathered in my study thus far, that language switch is particularly likely when the NL and TL have close cognates (as, with Turkish "balon" and English "balloon")—or, when the appropriate NL term approximates the desired TL *form and/or content* in some way. For example, all three Turkish subjects in my study referred to a caterpillar as a turtle. The Turkish word for catepillar is "tirtil". One can only speculate about what is going on here. It would appear that when the Turkish subjects initially used the term "tirtil", the experimentor did not reject it outright; she may have evidenced recognition of the word in some involuntary way which encouraged the subjects to continue using the term, even though it is clearly inappropriate. Something in the nonverbal interaction between subject and experimentor may have been at work to encourage language switch in these cases. Compare the following bits of data produced by B.L., the first of which shows language switch and the second of which does not:

a) B.L.: There are a little animal on the mushroom but I don't know what's its name. It's turtle? I think it's turtle . . .

b) B.L.: We call it *nargile*. What is this? What called?

In the first example, B.L. tests the NL term and decides to use it; in the second example, she tests the equivalent NL term and rejects it.

(4) *Appeal for assistance* ("appeal to authority" in TCD) occurs when the learner asks for the correct term: asks the experimentor, any native speaker, or even refers to a dictionary. For example, B.L. when confronted with the waterpipe above, says, "What is this? What called?"

(5) *Mime* refers to the use of nonverbal communication strategies by a second-language learner. For example, one subject in my study said, "and everybody say (claps his hands)". That is, it is possible to act out an action in order to communicate; in fact, this is one of the most basic ploys used when one is aware of one's lack of crucial TL vocabulary.

It may be helpful in gaining an understanding of the way these categories operate to examine what was probably the most difficult object to describe in English—the waterpipe in the Arthur Rackham illustration of the Alice in Wonderland caterpillar. In fact, this picture frame was atypical in that it elicited communication strategies in the subjects' NL descriptions, as well as in their IL descriptions, since some of the objects depicted are rather unusual. The waterpipe turned out to be difficult in both NL and in English (see Table 2). C.T. uses topic avoidance in both NL and IL, refusing to comment on the waterpipe at all. R.D. uses a strategy of approximation, calling the waterpipe a "lamp" in both NL and IL; he clearly knows the term is inappropriate, since in NL he describes the caterpillar as "smoking" from the "lamp". G.U. initially uses topic avoidance in English; when asked later to describe the object, he uses circumlocution and finally approximation ("pipe"). B.L. uses language switch and then appeals for assistance. D.R. uses a form of approximation ("pipe . . . water"). A.H. uses circumlocution in English to describe the object; interestingly

TABLE 2

The Waterpipe Responses

Learner	L1	Interlanguage	Native Language
C.T.	Spanish	ø	ø
R.D.	Spanish	some lamp	ø. (later) una lampara
G.U.	Spanish	ø. (later: This is a bot, bottle, with a . . . no, this is not wire . . . (E: a tube) a tube. I don't know exactly the name . . . pipe.	una pipa (a pipe)
B.L.	Turkish	We call it nargile. What is this? What called?	nargile (waterpipe)
D.R.	Turkish	pipe, water	nargile (waterpipe)
A.H.	Turkish	She is, uh, smoking something, I don't know what's its name. That's uh, Persian, and we use in Turkey, a lot of.	Yaninda bir . . . oh, I forgot the Turkish (laugh) var. (Next to him there is a . . . .)
M.R.	Mandarin	The witch is smoking, in a strange way.	Neichouyen de yangdz hen (that smoking form very) ougwai. Neige yen tung (strange that smoke pot) ye hen gugwai (also very strange)
J.O.	Mandarin	ø. (E: Do you know the name of this?) ø (E: It's a waterpipe) A waterpipe. Jyau shwei yen. (Called water pipe)	ø
M.S.	Mandarin	A very strange machine and some smoke went out of that bottle, I mean the originally machine. (later) E: What is this? M: Originally I forgot. I add something, I say just like smoke and pipe and connect with bottle. E: Do you have a single word in Mandarin that describes this? M: No. Uh, yes um we, maybe we have one, ju, just like, uh, do you know, um a, a poison there is, uh, no E: A drug? *Opium?* M: *Yes* yes. And, uh, do you know the old, uh, habit in, in old China? E: Yeah, yeah. Smoking *opium*. M: *Yeah*, smoking E: Opium.	Jeige dungsy (?) gwandz. (this thing pipe) Jeige gwandz lyanje pingdz (this pipe connect bottle) Jemma yichwei jyou mausye (this way blowing smoke comes out) mauyen chulai hausyang chouyen yiyang. (smoke comes out the same as smoking)

TABLE 2—*(Continued)*

M: Op . . .
E: Opium.
M: How do you spell?
E: O-P-I-U-M.
M: O-P-I-U-M. Is a . . .
E: It's a *drug*.
M: *Is a* kind of, plant?
E: Mmhm. It's a *poppy*.
M: *Opium*.
E: It's a poppy plant that grows and
the flower is *very bright*.
M: *Yes, yes*, oh.
E: Opium.
M: Oh. Yes, we, we have one called . . .
Mandarin is ye pien yen.
(opium pipe)

enough, he forgets the correct term in Turkish and resorts to message abandonment in his NL. M.R. uses circumlocution to describe the action in English. J.O. uses topic avoidance in both English and Mandarin; when asked later for the English word, she again does not respond. Interestingly, when she learns the correct English term, she translates it literally into Mandarin, using conscious transfer in reverse. M.S. uses the strategy of circumlocution initially to describe the waterpipe in both English and Mandarin. The role of the communication strategy of paraphrase in the negotiation of meaning is perhaps best observed in the exchange between M.S. and the experimenter which is recorded at the bottom of Table 2. The communication strategy being used is circumlocution as the learner names a series of items related to opium and hence to waterpipes; the function of the strategy seems to be to engage both M.S. and the experimentor in a joint negotiation of meaning which resembles a game of twenty questions.

A preliminary examination of the data shows definite differences in the learners' overall approach to the storytelling task. C.T. and R.D., whose proficiency levels appear to be very similar, and both of whom speak Spanish as a NL, perform very differently in this regard. C.T. tells her stories quickly in both NL and TL, and is content not to provide much detail. In contrast, R.D. elaborates on the pictures at great length in both NL and IL, with frequent appeals for assistance in IL and commentaries on his own progress (such as, "God, will I ever learn English?"). R.D. seems to use his imagination more, creating names for the characters and dwelling on the dramatic development of the storyline. In view of these individual differences in the sheer volume and nature of the data produced by the same task, I believe it is doubly important to elicit data in *both* NL and IL in order to avoid the error of assuming that C.T., for instance, is using an avoidance strategy simply because she produces less data than R.D. We can see that C.T. also produces less data in NL than does R.D.—and we clearly cannot assume that she does so because she is avoiding NL forms or structures.

A sampling of the learners' performance in response to the seven target concepts considered for the purpose of this report, was used to determine whether individual learners did appear to prefer certain types of strategies over others. As I mentioned before seven different target concepts were chosen, and the strategies used by each subject to communicate those concepts were tabulated. Table 3 summarizes the preferences of each subject in terms of the five major categories of communication strategy.

The subjects are listed in Table 3 in what I estimated to be a rough order of proficiency to English, with M.S. at the top appearing to be most proficient, and J.O. at the bottom as least proficient. It is clear that, even on the basis of this restricted body of data, the learners do seem to exhibit decided preferences for certain types of strategies and not for others. When M.S. does not use the correct TL form, he seems to rely quite heavily on paraphrase. J.O., on the other hand, uses avoidance strategies much more extensively. B.L. (third from the bottom) relies quite heavily on conscious transfer and appeals for assistance.

TABLE 3

Strategy Preferences, Evidenced in Sampling of Seven Target Concepts

	Strategy 1	Strategy 2	Strategy 3	Strategy 4
M.S.	0	5	0	0
A.H.	0	3	2	1
G.U.	1	7	0	1
D.R.	0	5	2	0
M.R.	0	4	2	4
R.D.	1	4	0	2
B.L.	0	2	4	3
C.T.	2	4	0	0
J.O.	4	2	1	1

Based on these preliminary patterns in the data, I would hypothesize that individual second-language learners do exhibit conscious communication strategy preferences. Further, my hypothesis is that a learner's first-language background in itself will not bias her towards any particular strategy preference. Rather, I believe that personality factors may correlate highly with strategy preference. That is, certain personality characteristics may be tied to a preference for avoidance strategies—and others to a preference for appeal to assistance. Further, I believe that strategy preference and second-language proficiency level may prove to be related, such that strategies of paraphrase could be increasingly preferred as successful second-language learners gain in proficiency.

My work in this area is just beginning. Clearly, these hypotheses must be tested from many different perspectives. But I believe investigation in this area of second-language acquisition will be exceedingly fruitful in the insights it will provide into the relationship of personality to second-language acquisition the nature of the relationship of language and thought, and perhaps even provide some insights into the nature of language proficiency.

REFERENCES

Adjemian, C. 1976. On the nature of interlanguage systems, *Language Learning 26*, 2, 297–320.

Gorbet, F. 1974. What the teacher can do. In *Errors: A New Perspective*, Ottawa, Canada, Public Service Commission of Canada, Research Division. Pp. 30-78.

Guiora, A. *et al. 1975*. Language and person—studies in language behavior. *Language Learning* 25, 1, 43–62.

Larsen-Freeman, D., and V. Strom. 1977. The construction of a second language index of development. Paper delivered at the First Annual UCLA Conference on Second Language Acquisition, February 1977.

Levenston, E. A. 1971. Over-indulgence and under-representation. In G. Nickel (ed.), *Papers in Contrastive Linguistics*, Cambridge, Cambridge University Press, 71–121.

Richards, J. 1975. Simplification: a strategy in the adult acquisition of a foreign language. *Language Learning* 25, 1, 115–126.

Selinker, L. 1972. Interlanguage. *IRAL 10*, 3, 209–231.

Tarone, E., A. Cohen, and G. Dumas. 1976. A closer look at some interlanguage terminology. *Working Papers in Bilingualism* No. 9, 76–90.

Tarone, E., U. Frauenfelder, and L. Selinker. 1976. Systematicity/variability and stability/instability in interlanguage systems. In D. Brown (ed.), *Papers in second language acquisition*, Ann Arbor, Language Learning.

Váradi, T. 1973. Strategies of target language learning communication. Paper presented at the VI Conference of the Romanian-English Linguistics Project in Timisoara, 1973.

The Nature of Linguistic Input in Formal Second Language Learning: Linguistic and Communicative Strategies in ESL Teachers' Classroom Language

Stephen J. Gaies
Indiana University

Compared to the amount of research which has investigated the processing strategies which learners use in first and subsequent language acquisition, very little attention has been paid to the nature and possible importance of language input data itself. This has been due, claims Landes (1975), to the emphasis placed by the nativist theory of language acquisition on children's innate capacities for language learning and the concomitant devaluation of the role of environmental (external) factors in language acquisition.

Recently, however, the need to investigate language input data more fully has been cited for both research in first language acquisition (Campbell and Wales, 1972; Clark, 1974; Landes, 1975) and studies of natural second language acquisition (Wagner-Gough and Hatch, 1975) and formal second language learning (Henzl, 1973, 1975; Gaies, 1976, a,b,c; 1977).

The study reported in this paper attempted to elucidate further the nature of linguistic input in formal second language learning situations. In the study, the classroom language of a group of teachers of adult ESL classes was investigated for evidence of linguistic and communicative strategies which might facilitate or otherwise influence the language learning task. Analysis of the data collected indicates that many linguistic and communicative strategies typically used by adults in their interactions with children in the early stages of language acquisition are equally characteristic of the classroom language of ESL teachers.

A recent and encouraging development in research in first and second language acquisition has been the renewal of interest in the role of the linguistic environment in language learning. For a number of years, many studies of language learning focused almost exclusively on the utterances which learners produced and paid little attention to the language directed at them. This neglect of language input data—this failure to pay attention to the language to which a learner is exposed—can be attributed in part to the methodological difficulties of such research, but even more so to the influence of the early nativist view of first and (as has been inferred by many) second language acquisition.

While the nativist theory of language acquisition continues to enjoy general acceptance among psycholinguists and applied linguists, it has been argued (Landes, 1975) that the early nativist view at least partially misrepresented the nature of language acquisition through its assumption that language learners are exposed to random, often ungrammatical, and essentially unstructured samples of language to be acquired. This notion led to the position espoused most notably by McNeill (1966) that the kind of linguistic input to which a learner is exposed is basically irrelevant, because it is the learner's innate pre-disposition to learn language that guides and shapes the acquisition process. In studies of first and second language learning, the overwhelming emphasis has in many cases been on examining the processes and strategies employed by learners as they progress toward a closer approximation of full target language competence. The abundance of "error analyses" of second language learners' utterances bears witness to the emphasis placed on learner strategies, to the primary focus on the language learner and the frequent accompanying neglect of the nature of the interaction of learner and linguistic environment.

Recently, however, the argument has been advanced (Campbell, 1972; Clark, 1974; Landes, 1975) that while the behaviorist claim that external, environmental sources are the sole determinants of learning is inadequate as an explanation of how language is learned, it is equally simplistic to ignore almost altogether the linguistic context in which language learning takes place. This argument is based on an accumulating body of empirical data which suggests the following points about the linguistic input to which first language learners are exposed:

(1) In terms of syntax, the language addressed to children by adults consists for the most part of consistently short, grammatical sentences. These sentences on the whole are transformationally simpler, involve far less subordination and self-embedding, and include a much greater number of interrogative and imperative sentences than those typical of adult-adult discourse.

(2) There is a tendency on the part of adults to use a relatively restricted lexicon when addressing young children. The token-type ratio, which is a measure of vocabulary diversity, of adult-child interactions has been found to be far smaller than that of samples of adult-adult discourse (Drach, 1969; Granowsky and Krossner, 1970).

(3) Characteristic phonological features of the language addressed to children by adults are a reduced rate of speech, clearer articulation, and an exaggeration of normal stress and intonation patterns.

These consistent adjustments in the language addressed to children by adults (and this includes parents, non-parents, and even older children) have been characterized by some researchers as constituting a special linguistic style called "motherese." It would thus be a situational register deemed appropriate for use when the interlocutor is a young child in the process of acquiring a language. It should be pointed out, however, that the degree of linguistic adjustment appears to be geared to either the changing age or the increasing skills of the child. That is, the older the child, the less radically the language addressed to him/her differs from that used between adults.

In addition to these linguistic modifications, adults employ a number of communicative and/or language training strategies in their verbal interactions with young children. One of the most salient of these strategies is *repetition*,[1] which is a recurrent technique thought to have potential accelerating effects on language acquisition. Snow (1972) has argued that repetition increases processing time, thus increases a child's chances of processing input. One study (Kobashigawa, 1969) of first language acquisition revealed that in a corpus of utterances addressed by a mother to her twenty-six month old son, 15% of the statements, 35% of the questions, and 60% of the imperatives were repeated with no long intervening pauses or activity.

Prodding characterizes instances when a parent (or other adult) makes it verbally clear that he/she wants the child to say or repeat something. Such verbal directions typically take the form of "Can you say . . . ?" or "Say . . .". A related device is *prompting*, in which, for example, a parent will show a child a picture of something which the child knows the word for and will ask, "What's this?" or "This is a *what*?" Prompting and prodding afford the child practice in using the language and are thought to have some effect on language acquisition, despite the fact that, as Slobin (1971) has pointed out, frequency of repetition sometimes weighs less heavily with children than the perception of an underlying pattern.

Another general strategy is *modeling*, which may involve something as simple as an adult's supplying the appropriate lexical item for a child who does not know the name of something in a picture. A more complex form of modeling takes place when an adult expands the child's utterances. These so-called "imitations in reverse," which are presumably performed by adults to check their understandings of children's utterances, have been thought to contribute to language development in that they recognize the truth value of a child's statements at the same time that they demonstrate to the child how those statements are encoded by his/her speech community at the very moment when the child is likely to be most attentive to such information. In contrast to responses to other forms of adult corrective feedback, children's imitations of adult expansions are frequently grammatically progressive. Empirical evidence for this was provided by Nelson et al. (1973), who found that syntactic development of an experimental group of forty-month-old children was significantly enhanced by their exposure to expansions which recast the subjects' sentences by providing a new syntactic structure.

Unfortunately, it is premature to make a precise statement of the ways in which the linguistic and communicative adjustments characteristic of the speech addressed to children facilitates or otherwise influences the language acquisition process. The issue is clouded by the fact that though the relative amount of adult linguistic input versus input from child-child interactions and adult linguistic performance heard by, but not specifically addressed to the learner, varies considerably from child to child within a culture and from culture to culture, the acquisition process proceeds at a remarkably uniform rate. At the very least, we have evidence that children incorporate either literally or through recombination sentences of units of speech even larger than sentences into their own linguistic repertoire from the language they hear around them, often without first internally analyzing them and deducing the underlying structure. Beyond that, the provisional assumption that the relative simplicity and organization of adult linguistic input and the use by adults of the training stategies just listed have some facilitating effect on the acquisition process, in that the child may more easily discover, for example, the basic sentence patterns, major constituent categories, and salient phonological features of the language, does not seem unreasonable.

---

[1] These categories are from Landes (1975).

Does the input to which formal second language learners are exposed through the oral classroom language of their teachers involve linguistic and communicative adjustments analogous to those which are characteristic of much of the adult input in first language acquisition? This was one question which the present study examined. The study was primarily an investigation of the syntax of the oral classroom language of eight ESL teacher-trainees enrolled in a Practicum course offered by the Program in Applied Linguistics at Indiana University.

Three of the subjects were highly proficient nonnative speakers of English who had had some experience teaching English in their home countries. The others were native speakers of English whose prior teaching experience was quite limited.

The subjects taught adult ESL classes as part of the Practicum course requirements. These classes, which are offered each semester, meet hourly four evenings a week for a period of ten weeks. In the Fall of 1975, when the data for the present study was collected, instruction was offered at four different levels. Each of these four levels was taught by two teachers, who shared the teaching responsibility equally. It should be mentioned that the subjects were given a great deal of freedom in making decisions about teaching materials and techniques (particularly the former, since no texts were used in these classes), curricular goals, and classroom management practices.

Each of the subjects agreed to let the researcher tape three of his/her classes: one each at the beginning, middle, and end of the ten-week period. In addition, the weekly meetings of the Practicum class, in which the subjects and their instructors discussed general and specific problems and approaches in teaching English to speakers of other languages, were taped so that samples of the language which the subjects used with each other—i.e., among linguistic peers—could be obtained.

From each of the twenty-four classroom tapes, a corpus was selected for analysis. In each case, the sample to be analyzed syntactically consisted of the first 500 words contained in utterances containing an independent clause spoken by the teacher during the actual class period. For the baseline language data collected in the Practicum class meetings, the first 500 words spoken in sentence-length utterances by a subject to the class as a whole and with the class' attention constituted the sample for that subject. For both the ESL classroom and baseline data, the non-sentential utterances—fragments, false starts, attention holders, etc.—which occurred in the corpus were tabulated separately.

The samples were transcribed and analyzed by the researcher. Syntactic analysis proceeded by segmenting the 500-word samples into T-units. A T-unit is defined as "one main clause plus any subordinate clause or nonclausal structure that is attached to or embedded in it" (Hunt, 1970, p. 4). This unit of syntactic analysis is objective and easy to compute, and in the last ten years it has gained increasing recognition as a far more valid index of syntactic complexity than other measures, including sentence length. A particularly attractive feature of the T-unit as an index of syntactic maturity (complexity) was revealed by O'Donnell, Norris, and Griffin (1967); in their study of syntax of the oral and written language of elementary schoolchildren, these researchers noted a close relationship between T-unit length and the number of sentence-combining transformations required to generate a T-unit. Differential mean length of T-unit, then, appears to reflect the relative degree to which users exploit the transformational resources of English.

Altogether, measures on six dependent variables were computed for each sample. These variables were: words per T-unit, ratio of clauses (main and subordinate) to T-units, words per clause, adjective (relative) clause per 100 T-units, adverbial clauses per 100 T-units, and noun clauses per 100 T-units.

Table 1 presents a comparison of the means of the subjects' classroom language and their speech among linguistic peers. The data as measured by all six dependent variables indicates an overall process of syntactic simplification in the classroom language. The subjects spoke in shorter clauses

TABLE 1

Comparison of Syntactic Complexity of Subjects' Baseline (Practicum Class Meetings) Language and Their Oral Classroom Language

	SOURCE OF LANGUAGE SAMPLE	
VARIABLE	BASELINE (N = 8)	ESL CLASSROOM (N = 24)
w/T	10.97	6.19
c/T	1.60	1.20
w/c	6.84	5.10
AD/100	11.59	2.54
AV/100	20.27	5.33
N/100	28.54	11.16

w/T = words per T-unit
c/T = clauses per T-unit
w/c = words per clause
AD/100 = adjective clauses per 100 T-units
AV/100 = adverb clauses per 100 T-units
N/100 = noun clauses per 100 T-units

and used fewer subordinate clauses per T-unit when addressing their students than they did when speaking to highly proficient interlocutors. Multivariate analysis of variance performed on the data revealed that the overall difference in syntactic complexity was highly statistically significant ($p = <.0001$).

The next step in the analysis involved the comparison of the oral performance of the subjects according to the level at which they were teaching. What emerges from the data (see Table 2) is an unmistakable and statistically significant ($p = <.0227$) relationship between the syntactic complexity of the subjects' classroom language and the level of proficiency of their students. At any level, the snytax of the teachers' oral classroom language is more complex than at the level immediately below it and less complex than at the level immediately above it; and *this is true for every one of the six criterion variables.*

TABLE 2

Comparison of Syntactic Complexity of Subjects' Oral Classroom Language by Level

	VARIABLE					
LEVEL	w/T	c/T	w/c	AD/100	AV/100	N/100
L_1 (Beginner) (N = 6)	4.30	1.02	4.20	0.00	0.76	1.60
L_2 (Upper Beginner) (N = 6)	5.75	1.14	5.04	1.46	3.64	6.92
L_3 (Intermediate) (N = 6)	6.45	1.24	5.18	2.26	8.40	13.54
L_4 (Advanced) (N = 6)	8.26	1.38	5.98	6.47	8.51	20.91
$\bar{x}_L$ (N = 24)	6.19	1.20	5.10	2.54	5.33	11.16

This syntactic adjustment evident in the subjects' oral classroom language is similar to that found both in previous studies (Henzl, 1973, 1975) of the verbal performance of foreign language teachers and in samples of adult language addressed to children. Equally similar was the way in which the rate of speech and vocabulary diversity of the subjects' classroom language increased as a function of the level of proficiency of the students they were teaching.

These similarities between adult input to children and the speech addressed to adult second language learners by their teachers are not suprising. After all, research by Ferguson (1975) and others has led to the claim that modifications of the full adult system of this sort are consistently made by members of a speech community on behalf of all learners of the language, whether or not the verbal interaction has explicit instructional goals. What is particularly noteworthy, however, is the presence in the oral classroom language of the subjects of both linguistic adjustments and the very training strategies characteristic of adult input to children.

Consider first the strategy of repetition. As mentioned earlier, repetition is an alternative or complement to linguistic simplification as a means of facilitating comprehension. It was for this reason that sentence-length utterances which were exact and immediate—that is, which occurred with no intervening pauses, activity, or change of interlocutor—repetitions were marked and tabulated separately in the analysis of the data. Repetition was used most frequently at the two lower levels of instruction, and practically not at all at the advanced level. An indication of how prevalent this strategy was for communication and teaching at the beginner level is the fact that utterances containing a total of 660 words were exact repetitions. In other words, more than 20% of the subjects' sentences were repeated in exactly the same form. Thus, at the beginner level, students not only heard input which was far less syntactically complex than that directed at more proficient students, but also had in general more time to process that input.

Prompting and prodding are strategies which can also be observed at the lower levels of formal classroom language instruction. These strategies, however, seem, even more so than repetition, to be very transitional devices, and their use is predictably restricted to those classroom activities which involve naming concrete objects—i.e., to the initial stages of instruction.

A more important classroom strategy is modeling, which was accomplished through two complementary but very different procedures. The first involved the use of fragments. If a subject asked a student a question and the student made an appropriate and syntactically complete response, the subject would often supplement the reinforcement of that response by repeating the crucial lexical or other item in the response. For example, if a student answered the question, "How did you come to class this evening?" by saying, "I came to class by bus," there was a tendency to respond, "Yes, by bus." Again, the strategy was most evident at the lower levels, which proceeded much more frequently than the more advanced levels through long series of short, teacher-initiated question-answer drills.

The alternative modeling technique—and the strategy which was most evident at all levels—was teachers' expansion of students' utterances. The stategy was often purely communicative—i.e., when it was used to check comprehension—but it served equally often, and again primarily at the lower levels, as a pedagogical strategy. The most obvious case of this was when a student responded to a question with a single word or phrase and the teacher expanded the response into a full independent clause.

What implications can reasonably be drawn from the similarity in linguistic adjustments and communicative/training strategies of the input data to which first and second language learners are exposed? First of all, the observed "training sessions" which result from the use of these strategies in a teacher's verbal interactions with his/her learners would be a major and identifiable criterion distinguishing formal from unstructured or "natural" second language learning, since in the latter these strategies would presumably be considered less frequently employed by speakers of a language whose contact with a learner of the language has no explicit pedagogical goals.

Again, there is additional support for Corder's (1967) position that the burden of proof remains with those who argue that first and second language learning proceed in a fundamentally different manner. This is not to suggest that there are no differences between first and subsequent language learning. One difference already alluded to is that many training strategies begin to be abandoned relative early—perhaps too early—in formal second language learning. And obviously, if we judge solely by the criterion of the relative percentage of instances on which first and second language learners attain full competence in the target language, there is a quantitative difference involved. However, just as a number of studies of second language learning have demonstrated that learners use processing strategies very similar to those employed by children acquiring their first language, we now have evidence, however preliminary, that the organization and presentation of the input with which they work is not unlike the primary linguistic input which children hear.

Another implication arises from the relationship between the perceived proficiency of the learners and the nature of their teachers' oral classroom language. Elsewhere (Gaies, 1976c), I have discussed the need for examining linguistic input as a factor in the development of transitional stages of competence or "interlanguages." Let me simply suggest here that as regards second language learning materials development, the principle that language drills should involve sentences which are increasingly longer and syntactically more complex (Steglitz, 1973) is indirectly validated by the findings of the present study. In view of the fact that teachers gear their own spoken language and communicative strategies to the proficiency of their students, the claim that grading a series of oral manipulation drills from easy to difficult by itself leads to an "artificial" exposure to the target language is not defensible.

A final and potentially the most immediately important implication of the relationship of the data from the present study to the linguistic input in first language learning is a revised notion of what is natural and what is artificial about formal (classroom) second language learning. Many of the kinds of foreign language classroom activities closely associated with the inductive or "audio-lingual" approach to language teaching, as well as many of the concomitant teaching procedures, are not at all different from the kinds of language training sessions which adult linguistic input provides children with. At the early stages of formal second language learning—at the period when students are traditionally provided the training in skills acquisition which is thought to be essential to more creative, sophisticated uses of the target language—adult language learners are asked to repeat (this is equivalent to the strategies of prompting and prodding), have items repeated for them to facilitate processing, are given verbal signals to imitate—this is equivalent to modeling—and often have their responses expanded by their teachers (which constitutes a reinforcement of the truth value of their response as well as a model of how that response is fully elaborated in the target language). There is nothing inherently artificial about these activities, at least if we judge by their role in first language acquisition. What makes them "unnatural," and perhaps more than anything else provokes dissatisfaction, is that these drills, these training strategies, are too often semantically divorced in every way from the classroom surroundings or the students' experience, past and present. Then they are unnatural and uninteresting. But this is no more necessarily the case in formal second language learning than it would be in first language acquisition.

REFERENCES

Campbell, R., and R. Wales. 1972. The study of language acquisition. In J. Lyons (ed.), *New horizons in linguistics*, Baltimore, Penguin. Pp. 242–60.

Clark, R. 1974. Performing without competence. *Journal of Child Language 1*, 1–10.

Corder, S. P. 1967. The significance of learners' errors. *IRAL 5*, 161–9.

Drach, K. 1969. The language of the parent: a pilot study. In The structure of linguistic input to children, Working Paper No. 14, Language Behavior Research Laboratory, University of California, Berkeley.

Ferguson, C. A. 1975. Toward a characterization of English foreigner talk. *Anthropological Linguistics 17*, 1–14.

Gaies, S. J. 1976a. The syntax of ESL teachers' classroom language: a preliminary report. Paper delivered at the Summer Conference on Second Language Learning and Teaching, SUNY (Oswego), July 16–18.

———. 1976b. ESL teachers' classroom speech: suport for the $L_1 = L_2$ hypothesis. Paper delivered at the Sixth Annual Conference of the New York State English to Speakers of Other Languages and Bilingual Education Association, Albany, October 22.

———. 1976c. Gradation in formal second language instruction as a factor in the development of interlanguage. *Lektos: Interdisciplinary Working Papers in Language Sciences* (Special Issue: Error Analysis and Language Testing, St. Clair, R., and B. Hartford, eds.), pp. 29–41.

———. 1977. Linguistic input in first and second language learning. Paper delivered at the Sixth Annual University of Wisconsin-Milwaukee Linguistics Symposium, March 18–19.

Granowsky, S. and W. Krossner. 1970. Kindergarten teachers as models for children's speech. *The Journal of Experimental Education 38*, 23–8.

Henzl, V. 1973. Linguistic register of foreign language instruction. *Language Learning 23*, 207–22.

———. 1975. Speech of foreign language teachers: a sociolinguistic analysis. Paper delivered at the Fourth International Congress of Applied Linguistics, Stuttgart, West Germany, August 25–30.

Hunt, K. W. 1970. Syntactic maturity in schoolchildren and adults. *Monographs for the Society for Research in Child Development* 35 (Serial No. 134).

Kobashigawa, B. 1969. Repetitions in a mother's speech to her child. In The structure of linguistic input to children, Working Paper No. 14, Language Behavior Research Laboratory, University of California, Berkeley.

Landes, J. 1975. Speech addressed to children: issues and characteristics of parental input. *Language Learning 25*, 355–79.

McNeill, D. 1966. Developmental psycholinguistics. In F. Smith and G. Miller (eds.), *The genesis of language*, Cambridge, Mass., MIT Press. Pp. 15–84.

Nelson, K. E., G. Carskaddon, and J. Bonvillian. 1973. Syntax acquisition: the impact of experimental variation in adult verbal interaction with children. *Child Development 44*, 497–504.

O'Donnell, R. C., R. C. Norris, and W. J. Griffin. 1967. *Syntax of kindergarten and elementary schoolchildren: a transformational analysis*. Research Report No. 8. Champaign-Urbana, Ill., National Council of Teachers of English.

Slobin, D. 1971. *Psycholinguistics*. Glenview, Ill., Scott, Foresman, and Co.

Snow, C. 1972. Mothers' speech to children learning language. *Child Development 43*, 549–65.

Stieglitz, F. 1973. *Teaching a second language: sentence length and syntax*. Research Report No. 14. Champaign-Urbana, Ill.: National Council of Teachers of English.

Wagner-Gough, J., and E. Hatch. 1975. The importance of input data in second language acquisition studies. *Language Learning 25*, 297–308.

Acquiring a Second Language in an Immersion Classroom

Sandra Plann[1]

University of California, Los Angeles

This paper is a report of an investigation of the acquisition of some Spanish morphemes by some native-English-speaking children who are acquiring their second language in an immersion classroom where Spanish is the sole medium of instruction. The investigation consisted of a cross-sectional test of the children in grades one through four, all of whom began the program in kindergarten. The objective was to study their acquisition of agreement rules for article-noun-adjective gender and number, and for verb person and number. While these agreement rules are not crucial for communication, they are extremely important if the children are to achieve native-speaker proficiency in Spanish.

No definite trend of improvement across grades was found in the acquisition of morphemes. With both the noun modifiers and the verbs, the children overused the forms of highest frequency, overextending them to the forms of lowest frequency. The systematic nature of their errors, along with their lack of definite improvement across grades, points to the development of a classroom dialect peculiar to Spanish Immersion students.

The reasons for their lack of progress are probably not cognitive, as these forms are mastered earlier by native-speakers. Instead, the persistence of these errors is attributed mainly to the large amount of incorrect peer input from their English-speaking classmates. To a lesser degree, peer group pressure may also argue against native-like mastery of these forms. It seems that children who acquire a second language in an immersion classroom develop a classroom dialect of the language.

Recent research has shown child second-language acquisition to be an orderly, predictable process, in many ways similar to first-language acquisition (Dulay and Burt, 1972, 1973, and 1974 a & b; Ervin-Tripp, 1974). However, these studies deal with second-language acquisition in a natural setting; that is, the child acquires his second language while living in an area in which the language is widely spoken natively. This situation is called "absolute immersion." Second-language acquisition in a "total immersion" situation, as is the case when a child is educated exclusively in a second language to which he has little or no exposure outside of the classroom, is the focus of this paper.

In a Culver City, California, elementary school in 1971 the Spanish Immersion Program (SIP) was begun. English-speaking kindergartners receive all of their classroom instruction in Spanish. The aim of the program is to

---

[1] My thanks to Evelyn Hatch for her comments and suggestions on this paper.

produce English-Spanish bilinguals. At the time of this study (February, 1976), ten children of the 1971 pilot group were in their fifth year of the Immersion Program, and a new group had entered the program each year; thus, there were five SIP levels, kindergarten through fourth grade. This program has been the object of considerable research. (See Campbell, 1972; Cathcart, 1972; Broadbent, 1973; Flores, 1973; Boyd, 1974; Lebach, 1974; Waldman, 1975; Plann, 1976.)

The linguistic aims of this program are that the children should, while maintaining "normal progress in the maturation process of their first language (English). . . . acquire native-like proficiency in speaking, understanding, reading and writing Spanish" (Campbell, 1972, pp. 87-88).

Research on the pilot group in both the second and third years of immersion demonstrated that their acquisition of the agreement rules for article-noun-adjective gender, as well as subject-verb person and number, was as yet incomplete (Flores, 1973; Boyd, 1974). For instance, "*yo va a la baño" and "*las niños quiere comprar los frutas" were typical utterances. While utterances such as these enable the children to communicate, a mastery of the agreement rules is essential if they are to achieve the goal of native-like proficiency in Spanish.

The purpose of this investigation was to conduct a cross-sectional study of the SIP children in grades one through four to measure their degree of mastery of article-noun and noun-adjective agreement, as well as subjective-verb person and number agreement in the present tense.

The Subjects

The subjects were twenty-four SIP children from grades one through four, all of whom entered the program in kindergarten. Three girls and three boys were selected randomly at each grade level.

Materials and Procedures

Due to the number of students involved in this study, elicitation methods were used exclusively to gather data. Four tasks were used: (1) a puzzle, (2) a story telling task, (3) a repetition task, and (4) a re-telling of the story used for the repetition task.

The testing was conducted between February 10 and 25, 1976. Each subject was tested on all tasks within a one-week period. All instructions were given in Spanish, and each child's performance was tape recorded.

A puzzle whose pieces were pictures of very familiar objects in various colors was used to test gender and number agreement of definite articles and adjectives with singular and plural masculine and feminine nouns. There were twenty-four pieces, twelve masculine and twelve feminine, half of which were singular and half of which were plural. Each child was given the puzzle frame and a sheet of paper with pictures of all the pieces. The child was instructed to ask the experimenter for each piece pictured on the paper and then place it in the puzzle. For example, the child had to ask for the yellow dogs ("Quiero los perros amarillos"), the orange turtle ("Quiero la tortuga anaranjada"), and so on.

A story telling task consisting of eight pictures was used to elicit relatively free speech data. The child was instructed to look at the pictures and tell the story.

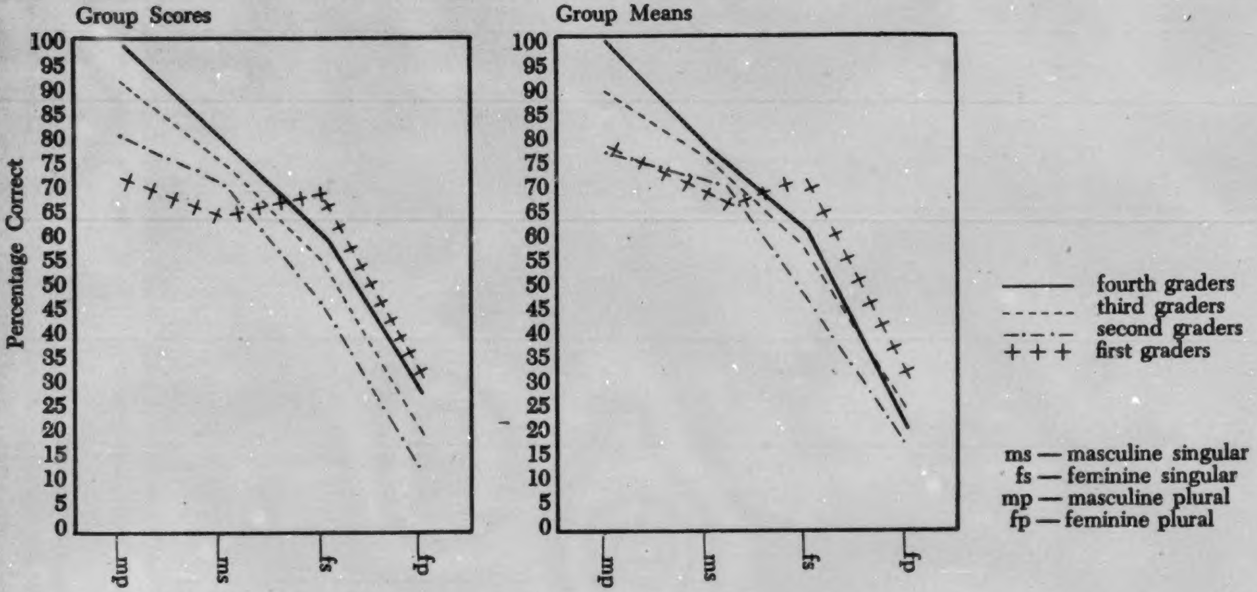
A repetition task in story format with accompanying pictures was included to insure the appearance of all relevant data and facilitate comparison across subjects.

The story re-telling task immediately followed the repetition task; the child was given the accompanying picture book and told to look at the pictures and tell, in his own words, the story he had just repeated.

A binary method was used in scoring each obligatory context; a morpheme was either correctly supplied (+1) or incorrectly supplied or missing (0). For each morphological form under consideration every child's score (the percentage of correct uses of the morpheme based on the number of obligatory contexts) on each of the four tasks was computed. The grade level averages were then found, according to two systems used by Dulay and Burt (1974b), the group score method and the group mean method. The group scores and the group means for each morpheme were ranked in decreasing order for each task and for each grade level to yield an accuracy order. To give an overall picture of the children's performance on each morpheme, two scores were also computed for each grade level based on the combined results of all four tasks. Thus, for each morpheme, an overall group score, the average of the four group scores on the four tasks, was found, as was an overall group mean. These overall scores were also ranked in decreasing order, to reveal a general picture of the accuracy sequence.

Findings

Definite Articles. In the findings on the four forms of the definite article a definite accuracy order was discernible for the children in grades two through four; this pattern, for both the group score and the group mean, is the same on each task as well as on the overall scores based on all four task scores combined (Fig. 1). Highest accuracy was attained on the masculine plural form ("los"), followed by the masculine singular ("el"); feminine singular ("la") was next, and feminine plural ("las") was lowest. The accuracy order of the first graders, unlike that of the children in the higher grades, varied from task to task; also, the first graders' accuracy order based on the combined scores of all four tasks differed from that of the children in the three higher grades. Their overall order was first masculine plural, next feminine singular, then masculine singular, and finally feminine plural. In comparing the combined scores of the three highest grade levels, a trend of improvement on each morpheme can be seen across grades, although, unfortunately, this trend could not be statistically tested since some children at each grade level managed to use fewer than the required three instances of each form.



Grade				
4	98%	80%	61%	28%
3	91%	75%	55%	20%
2	80%	71%	47%	13%
1	72%	64%	68%	32%

Grade				
4	100%	78%	61%	21%
3	89%	76%	57%	24%
2	77%	69%	47%	17%
1	78%	66%	71%	32%

Figure 1. Comparison of Overall Definite Article Sequences: All Grades.

Across all four grade levels, the only form to reach the traditional level of mastery, 90%, was the masculine plural. On the group score and the group mean, the fourth graders achieved 98% and 100% accuracy, respectively, and the third graders 91% and 89%, respectively. These were the only groups to reach this level, and on both the overall group scores and group means the masculine plural was the only definite article to reach the 90% level for any grade.

Indefinite Articles. In the use of the masculine and feminine singular forms of the indefinite article ("un" and "una"), the masculine was the form of highest accuracy at all grade levels. On all tasks except the third graders' story retelling task, it exceeded the 90% acquisition criterion. On the overall scores (Fig. 2) the fourth graders scored lowest on the masculine form by a few points but highest on the feminine form. The feminine form never approached the 90% accuracy level.

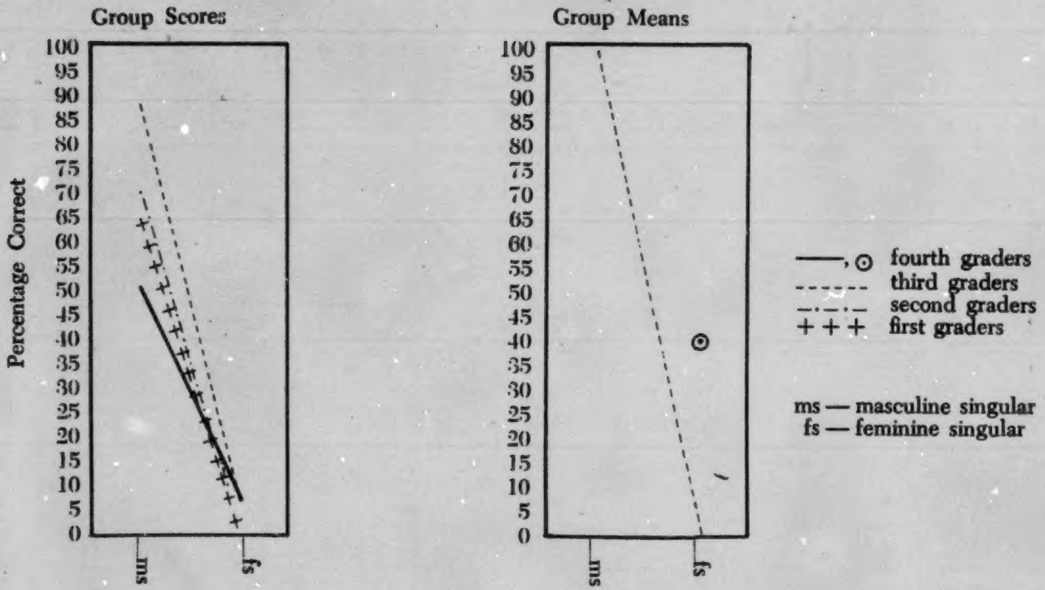
Adjectives. For the four adjective forms on all but the second graders' repetition and retelling tasks, the children had highest accuracy on the masculine singular form. With only two exceptions (the first and fourth graders' repetition task), the two masculine forms were always of highest accuracy and the two feminine forms of lowest accuracy. The overall accuracy order for grades one, two and four was masculine singular, masculine plural, feminine singular, and finally, feminine plural, while the feminine singular ranked lowest for the third graders, below feminine plural (Fig. 3).

In comparing across grades, for the masculine singular form the fourth graders scored highest, followed by the third graders; the overall group score for the first graders was slightly higher than that of the second graders. (Insufficient data made computation of a group mean for the first graders impossible.) With the masculine plural form, the second graders' overall group score was the highest, followed by the fourth graders, then the first graders, and lastly the third graders, although these differences were not statistically significant on the puzzle task when an ANOVA was run. For the two feminine forms, there was a general trend of improvement across grades, although in the one task where it was possible to run a statistical evaluation, the difference across grades was not significant.

For the masculine singular form, both the overall group score and the overall group mean of the third and fourth graders surpassed the 90% acquisition criterion. While the second graders' overall scores were below 90%, the first graders' overall group score was 90%. Performance on all other forms by all groups was below 90%.

Verbs. In the findings on the present tense verb forms there was much variation, both in percentage scores and in sequences across grades and across tasks. The only pattern of accuracy common to all tasks and all grades was the consistently highest accuracy of the third person singular form. Also, there was an increase in the correct usage of this form across grade levels (Fig. 4). In the overall totals, none of the forms reached the 90% acquisition criterion.

An error analysis revealed certain regularities. With the definite articles,

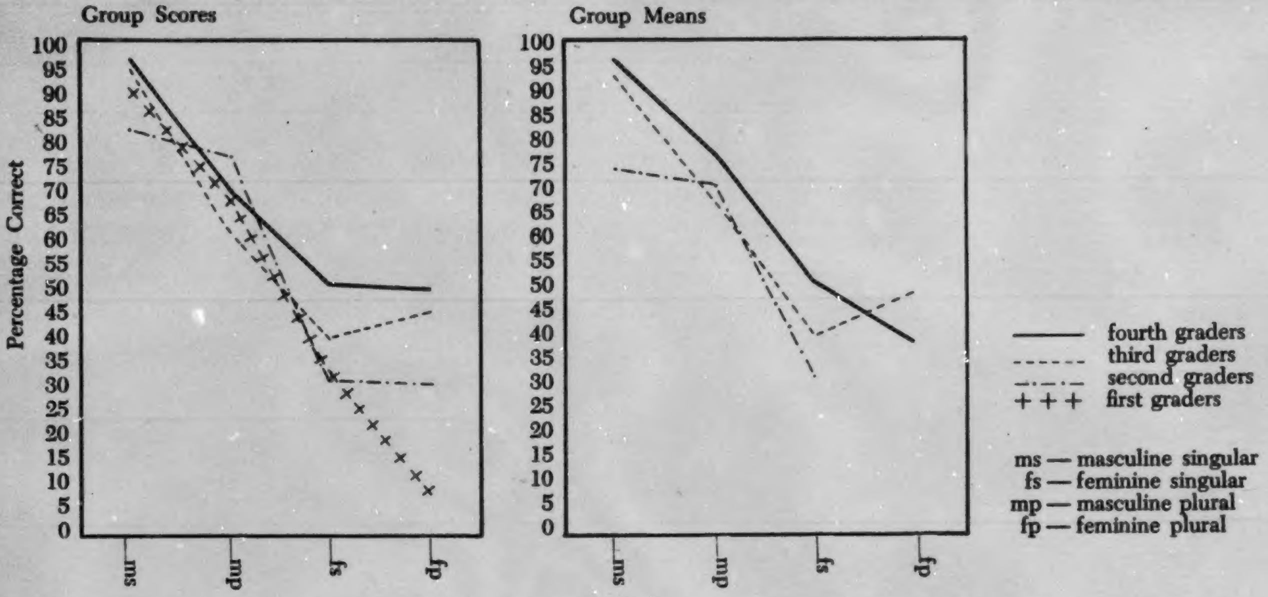


Grade		
4	91%	51%
3	93%	13%
2	100%	30%
1	100%	34%

Grade		
4	--*	42%
3	100%	2%
2	--*	--*
1	--*	--*

*insufficient data to compute group mean

Figure 2. Comparison of Overall Indefinite Article Sequences: Story telling, Repetition and Story Retelling Tasks, All Grades.

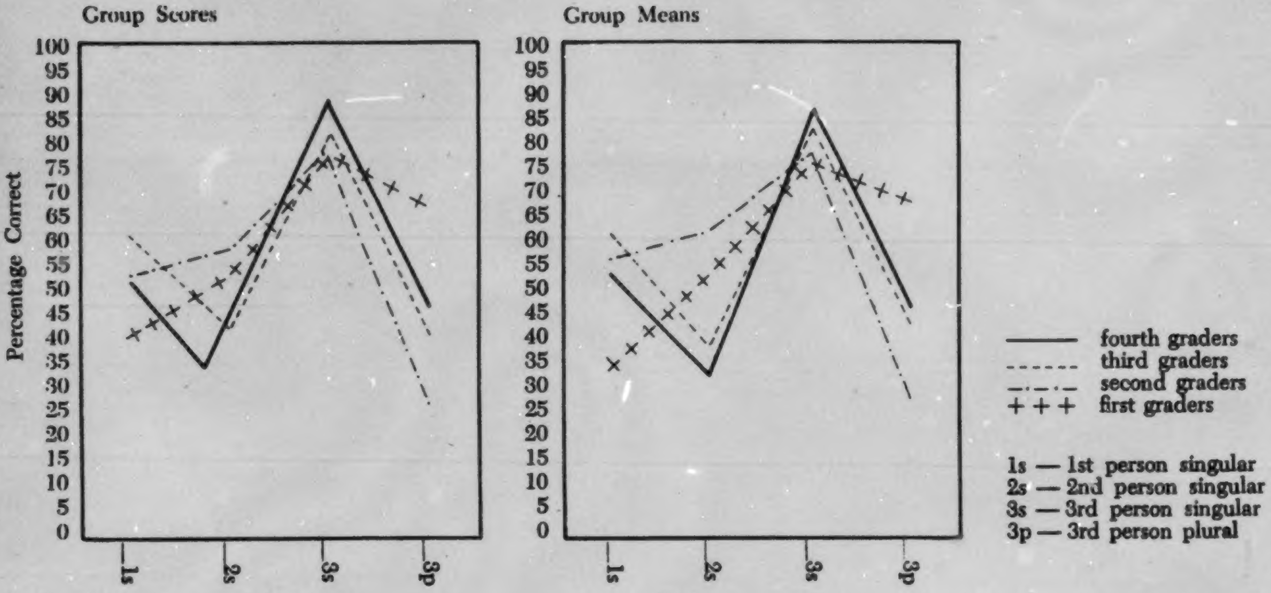


Grade				
4	96%	69%	51%	50%
3	94%	61%	40%	45%
2	82%	77%	32%	31%
1	90%	68%	33%	9%

Grade				
4	97%	78%	51%	39%
3	93%	68%	39%	48%
2	74%	71%	33%	--*
1	--*	--*	39%	--*

*insufficient data to compute group mean

Figure 3. Comparison of Overall Adjective Sequences: All Grades.



Grade				
4	52%	34%	88%	47%
3	61%	42%	81%	42%
2	53%	58%	77%	28%
1	40%	53%	77%	66%

Fig. 20-a

Grade				
4	52%	32%	85%	46%
3	59%	37%	81%	42%
2	54%	60%	77%	28%
1	32%	52%	74%	67%

Fig. 20-b

Figure 4. Comparison of Overall Adjective Sequences: All Grades.

errors were generally between masculine and feminine pairs ("el"/"la" and "los"/"las"); thus, the most accurate definite article, the masculine plural, was the form most often substituted for its feminine counterpart, the least accurate form. The children alternated in their use of the two forms of intermediate accuracy, the masculine and feminine singular. Thus, while errors like "*el pelota" and "*las niños" were common, errors like "*la pelotas" and "*los niño" were virtually non-existant. These errors indicate an acquisition of number, but not of gender agreement rules.

With the two singular indefinite articles, the form of higher accuracy was the masculine ("un"); the highest accuracy on the feminine form ("una") was attained by the fourth graders, who correspondingly scored lowest on the masculine form. This might mean that the fourth graders are making progress in their acquisition of the feminine form, and that it is temporarily displacing the overextended masculine form.

With the adjectives, errors with the singular forms were between masculine and feminine pairs, for instance "*una pelota amarillo," or "*una casa nuevo." In errors involving the plurals, the masculine singular (the form of highest accuracy) was most often substituted for the masculine plural ("*dos conejos blanco"), while the masculine plural (the form of second highest accuracy) was most frequently substituted for the feminine plural ("*tres mariposas amarillos"). These errors, too, suggest a better mastery of number than of gender.

With the verbs, the form of highest accuracy, the third person singular, was also the form most frequently overextended; utterances like "*ellos dice adíos," and "*yo quiere una" were very common.

Discussion

The test results, then, indicated that the children were still at an intermediate stage in their application of these morphological agreement rules. In no case were both members of a pair firmly established; rather, high accuracy of one form was usually accompanied by its overextension to the corresponding form which was, consequently, of low accuracy. Further, the children alternated frequently in their use of forms of intermediate accuracy. As Melanija Mikes (1967) pointed out, a category can be considered acquired only if the correlating forms are also correctly used. For instance, at all grades the children achieved high accuracy levels of the masculine indefinite article, "un," but it cannot be considered established because of the low accuracy levels of the feminine form, "una." Furthermore, Andrew Cohen (1974) found that the acquisition process of articles and adjectives, for at lease some SIP students, consisted of first the exclusive use of one form, and later, upon becoming aware of the other corresponding form, an alternation between the two or even an overextension of the second form. It is therefore doubtful that any of these forms, even those of very high accuracy, can be considered acquired.

While the SIP children have not as yet mastered these agreement rules, errors are an integral part of the learning process, and a comparison of the SIP students with other children acquiring Spanish demonstrates that difficulties with these morphemes are common to many young Spanish-speakers. Several researchers (Gonzalez, 1970; Brisk, 1973; Mazeika, 1973; Fantini, 1974; Cohen, 1975a) observed similar errors in the acquisition of Spanish by native Spanish-English bilinguals in the United States, as did Dato (1970) in the acquisition of Spanish by English-speaking children living in Madrid. However, while these are clearly developmental errors, they are more typical of younger language learners; with children the ages of the SIP students, such errors are either eradicated or much less frequent. In light of the progress of other children acquiring Spanish, lack of cognitive development does not adequately explain the SIP children's low accuracy on these forms; this suggests that factors other than cognitive maturity are involved.

A learner's speech inevitably deviates from that of a native speaker; his interium linguistic system, or "interlanguage" (Selinker, 1972, 1975), represents the learner's attempt to reproduce the target language. Characteristics of an interlanguage are its mutual intelligibility among speakers, its systematicity at one particular point in time, the stability of errors, and backsliding or reappearance of errors thought to be eradicated. Selinker (1972) pointed out that interlanguage errors are subject to fossilization:

> Fossilizable linguistic phenomena are linguistic items, rules, and subsystems which speakers of a particular native language will tend to keep in their interlanguage relative to a particular target language, no matter what the age of the learner or amount of explanation and instruction he receives in the target language. (p. 215)

This description of an interlanguage is applicable to the SIP children, particularly those who have been in the program for several years. The lack of a definite trend of improvement across grades suggests that these morphological errors have fossilized. Furthermore, they are not amenable to correction. The third grade teacher has tried to call the Children's attention to these errors by orally correcting them and having them repeat the correction; the fourth grade teacher has attempted to teach these grammar points more formally, giving the class explanations at the blackboard and then having the children do oral drills and written follow-up exercises. However, although the students seem to grasp the concepts, both teachers admit there has been little improvement in the children's speech. Clearly, these persistent errors have fossilized.

Language learning is a complex process, related both to the child's cognitive abilities and his interaction with his environment. Some of the factors relevant to second language acquisition, or non-acquisition, include the learner's attitudes and motivation, his linguistic needs, and the circumstances of his exposure to the target language. In acquiring a second language, a learner's affective orientation towards the language and its speakers can greatly influence his learning achievement. Lambert noted that "the whole process of becoming bilingual can be expected to involve major conflicts of values and alliances"

(1967, p. 91). The Spanish language in California has less social status than English, and the children are doubtlessly, to some degree, aware of its inferior status. If a learner has no desire for greater identification with the speakers of the target language, his motivation to improve will not be strong (Nida, 1971). While the children have a favorable orientation towards Spanish and the Spanish-speaking community (Cathcart, 1972; Lebach, 1974; Waldman, 1975), their attitudes towards perfecting their command of Spanish may be influenced by the language's lower prestige.

The needs of a language learner also affect his degree of mastery of a language. The emphasis in the Culver City classrooms has always been on communication, and the children's Spanish is sufficient to express themselves and understand their teacher. Jain (as quoted in Taylor, 1974) wrote that a student's interlanguage is not transitional but rather represents a functional competence; a learner, having learned enough to fulfill his needs, will stop learning, and will thereafter always speak an interlanguage with fossils. For their classroom needs, the Spanish of the SIP students is adequate; any further improvement at a morphological level may seem unnecessary.

A third consideration in language acquisition are the circumstances of the learner's exposure to the target language. The exposure time of the SIP children is about five hours a day in school. Outside of the classroom, the children rarely use their Spanish nor have any contact with the Spanish-speaking community. Native-speaker models consist of the teacher, occasionally a bilingual aide, and a few Spanish-speaking peers. However, these people account for only a small part of their linguistic input; the majority of their input is from native-English-speaking classmates, themselves learners of the target language. Consequently, the majority of their input, which is considered to be of primary importance in morpheme acquisition, consists of incorrect peer language. (See Larsen-Freeman, 1976 a & b.) A related factor is peer pressure. Within the SIP classroom the children probably want to speak Spanish like the majority of their classmates, the other English-speakers. These two factors, primarily the large amount of incorrect peer input, and possibly peer group pressure to talk like the other English-speaking students are, I believe, the main causes of the SIP children's low accuracy on these morphemes.

These problems are inherent in an immersion situation, and combined with limited linguistic needs and, possibly, the lower status of Spanish in the community, they argue against the SIP children perfecting their command of their second language. Instead, the end result is that the children develop and reinforce their own classroom dialect. The fossilization of forms, particularly at the morphological level where semantic power is low, is perhaps an inevitable by-product of acquiring a second language in an immersion classroom.

A comparison of the progress of the SIP children with that of the students in the St. Lambert French Immersion Program (FIP) provides another example of the emergence of a classroom dialect among immersion students. There are many similarities between the two immersion programs. While the FIP children have positive attitudes towards French Canadians (Lambert and Tucker, 1972), French is nonetheless the language of a minority group and has less prestige than English (Lambert, 1967). The FIP students, outside of school, seldom use their French and have little or no contact with members of the French-speaking community (Selinker, 1975). These children, like the Culver City students, have made remarkable progress in their second language (see Lambert and Tucker, 1972); however, the FIP children have many grammatical difficulties in common with their SIP peers. They make more errors than native-speakers in their French production, particularly errors of gender (Lambert and Tucker, 1972). With the verbs, the children pick one form (the infinitive) and overextend it to all cases (Selinker, 1975). Thus, the FIP students, too, have developed a systematic interlanguage, complete with fossilizations of forms of low semantic value. Their difficulties with gender and strategies of overgeneralization are similar to those of their SIP peers.

These findings seemingly contradict those of other second language acquisition studies which found that the errors made by child language-learners were developmental and eradicated over time (Dulay and Burt, 1973; Ervin-Tripp, 1974). However, those subjects were acquiring their second language in natural situations and in the presence of a variety of native-speaking peers. This may be the crucial difference in the two situations; in an immersion classroom, without sufficient contact with peer native-speakers, children's errors seem to fossilize.

REFERENCES

Boyd, Patricia. 1974. Second language learning: the grammatical development of anglo children learning through Spanish. Unpublished master's thesis, University of California, Los Angeles.

Brisk, María Estela. 1974. New Mexican Spanish syntax of the five year old. In Garland D. Bills (ed.), *Southwest area linguistics*, San Diego, California, Institute for Cultural Pluralism, School of Education, San Diego State University.

Broadbent, Robert Redd. 1973. Some consequences of following an elementary school curriculum in a second language. Unpublished master's thesis, University of California, Los Angeles.

Campbell, Russell N. 1972. Bilingual education in Culver City. *Work Papers in Teaching English as a Second Language* 6, 87–92.

Cathcart, Ruth. 1972. Report on a group of Anglo children after one year of immersion in instruction in Spanish. Unpublished master's thesis, University of California, Los Angeles.

Cohen, Andrew D. 1974. The Culver City immersion program: how does summer recess affect Spanish speaking ability? *Language Learning* 24, 53–68.

———. 1975a. *A sociolinguistic approach to bilingual education: experiments in the Southwest*. Rowley, Massachusetts, Newbury House, Inc.

———. 1975b. Forgetting a second language. *Language Learning* 25, 127–138.

Dato, Daniel P. 1970. The development of the Spanish verb phrase in children's second language learning. In Paul Pimsleur and Terence Quinn (eds.), *The psychology of second language learning: papers from the second international congress of applied linguistics*, Cambridge, The University Press.

Dulay, Heidi C., and Marina K. Burt. 1972. Goofing: an indicator of children's second language learning strategies. *Language Learning* 22, 235–252.

———. 1973. Should we teach children syntax? *Language Learning* 23, 245–258.

———. 1974a. A new perspective on the creative construction process in child second language acquisition. *Language Learning 24*, 253–278.

———. 1974b. Natural sequences in child second language acquisition. *Working Papers in Bilingualism* 3, 44–67.

Ervin-Tripp, Susan. 1974. Is second language learning like the first? *TESOL Quarterly 8*, 111–127.

Fantini, Alvino Edward. 1974. Language acquisition of a bilingual child: a sociolinguistic perspective (to age five). Ph.D. Dissertation, University of Texas, Austin.

Flores, Marco. 1973. An early stage in the acquisition of Spanish morphology by a group of English-speaking children. Unpublished master's thesis, University of California, Los Angeles.

Gonzalez, Gustavo. 1970. The acquisition of Spanish grammar by native Spanish speakers..Ph.D. dissertation, University of Texas, Austin.

Lambert, Wallace E. 1967. A social psychology of bilingualism. *Journal of Social Issues 23*, 91–109.

——— and Richard Tucker. 1972. *Bilingual education of children: the St. Lambert experiment*. Rowley, Massachusetts, Newbury House.

Larsen-Freeman, Diane. 1976a. An explanation for the morpheme acquisition order of second language learners. *Language Learning 26*, 1, 125–134.

———. 1976b. ESL teacher speech as input to the ESL learner. *Work Papers in Teaching English as a Second Language 10*.

Lebach, Susan M. 1974. Report on the Culver City immersion program. Unpublished master's thesis, University of California, Los Angeles.

Mazeika, Edward J. 1973. A comparison of a monolingual and a bilingual (Spanish-English) child. Paper presented at the Biennial Meeting of the Society for Research in Child Development at Philadelphia, Pennsylvania, March 29—April 1.

Mikes, Melanija. 1967. Acquisition des catégories grammaticales dan le langage de l'enfant. *Bulletin de l'Institut National d'Orientation Professionnelle 5*, 22–30.

Nida, Eugene A. 1971. Some psychological problems in language mastery and retention. In Paul Pimsleur and Terence Quinn (eds.), *The psychology of second language learning*, edited by Cambridge, The University Press.

Plann, Sandra. 1976. The Spanish immersion program: towards native-like proficiency in a classroom dialect?" Unpublished master's thesis, University of California, Los Angeles.

Selinker, Larry. 1972. Interlanguage. *IRAL 10*, 209–231.

———, Merrill Swain and Guy Dumas. The interlanguage hypothesis extended to children." *Language Learning 25*, 139–152.

Taylor, Barry P. 1974. Toward a theory of language acquisition. *Language Learning 24*, 23–35.

Waldman, Ellyn Louise. 1975. Cross-ethnic attitudes of Anglo students in Spanish immersion, bilingual, and English schooling. Unpublished master's thesis, University of California, Los Angeles.

The Relationship Between Self Esteem and the Oral Production of a Second Language

Adelaide W. Heyde[1]

English Language Institute
University of Michigan

The first part of this article focuses on a review of the literature and theoretical aspects of self-esteem including definitions, testing, and research relating self-esteem to learning. The second part proposes a study using two groups of subjects (one learning English; the other, French/Spanish). The research will explore the relationship between the three levels of self-esteem (global, specific, and task) and the oral production of a second language. It is hypothesized that: (a) Global self-esteem (GSE) scores, specific self-esteem (SpSE) scores, and task self-esteem (TSE) scores will intercorrelate positively, with TSE scores most closely correlated with performance on an extemporaneous speech; (b) subjects with high self-esteem (HSE) will receive significantly higher scores from their teachers on their oral performance than subjects with moderate or low self-esteem, and HSE subjects will rate themselves higher on their oral performance; (c) SpSE measures will have a stronger relationship than GSE measures to predict performance on the extemporaneous speech. Samples of global, specific, and task measures are included. The paper contains a brief summary of a pilot study conducted in April, 1977. The conclusion discusses expected results, limitations of the study, and suggestions for further research.

For some time now, linguists, psycholinguists, and psychologists have been in search of the characteristics which comprise the "Good" or in some cases the "Bad" language learner. Researchers have tagged his speech steps through contrastive analysis, interlanguage, and a variety of affective variables such as motivation, empathy, field independence, integrative motivation, attitudes, and others. Although no complete and accurate explanation exists of how a person learns a second language, we know that a variety of affective and cognitive factors interact to effect the acquisition, the use, and the output of a second language. Errors in second language output are attributable to one of three factors: 1. the absence of the rule in the learner's cognition; 2. the presence of the rule in the learner's cognition and failure to apply it due to a non-affective factor (i.e. incomplete information about the rule and its use, insufficient time

---

[1] I wish to thank H. Douglas Brown, W. Thomas Schomaker, John A. Upshur, and the Good Language Learner Research Group for their support and encouragement in my undertaking of this project.

to process and use the rule, etc.); 3. the presence of the rule in the learner's cognition and failure to use it due to interference from one or more affective states. The present study focuses on an exploration of the third factor—affective variables, particularly, one type of attitude, and its relationship to second language output.

Attitudes can be classified into two categories: external and internal. Attitudes, as Lambert and Gardner (1972) use the term, are external in nature. External attitudes refer to those evaluations which the learner directs towards factors outside himself (e.g. the worth of the second language, of the second language's culture, of his native language, of his native culture, etc.) Internal attitudes, on the other hand, refer to those evaluations which the learner directs towards factors inside himself (i.e. his competency, his ability, his self-worth). Within this study, another term for internal attitudes is self-concept or self-esteem. Self-concept is the sum of one's attitudes towards himself—the evaluation of what a person feels he is.

Both internal and external attitudes effect motivation and behavior. Research by Gardner and Lambert (1972) shows the relationship among external attitudes, motivation, and second language behavior. Their term, "integrative motivation", refers to the learner's desire to be a part of the culture of the second language community and his positive attitude toward the people and language of the second culture. As for internal attitudes and their relationship to motivation and behavior, Barksdale (1972) suggests that self-esteem is a precursor to motivation (i.e. I must feel confident about my ability to survive a change in my life in order to risk the process of changing my behavior, personality, language behavior, etc.) Furthermore, Bills (1951) suggests that our desire to maintain and enhance the consistency within our value system (self-concept) motivates our behavior. In addition, Miskimins (1973, p. 32) suggests that what a person does "depends on his self concept. Thus an adequate account of motivation requires getting at the manner in which persons treat themselves." If we acquire the view that we are undesirable, worthless, or bad, we tend to act accordingly (Fitts, 1965, p. 1); or if we feel that we can't speak a second language, or if we feel that we don't have the ability to learn a second language, our language behavior will act accordingly. In short, the level of self-esteem may either block or facilitate the cognitive domain in its application of the rules and thereby affect the second language output.

The relationship of self-esteem to second language behavior can be described using Krashen's (1971) Monitor Model. Krashen distinguishes between acquisition and learning by describing acquisition as a subconscious process and learning as conscious process occuring as the result of a formal learning situation or self study (1977, p. 153). His diagram (1977, p. 154) shows the acquisition process occurring subconsciously before learning. Learning, acting as the Monitor, is the conscious application of acquired rules to second language output before speech occurs.

The model, he claims, explains the discrepancies found in research between attitudes and aptitude. Attitudes (p. 158) are subconscious and related directly to acquisition. Aptitude, however, relates to Monitor competence which is why the effects of aptitude appear in tests which encourage monitoring. The model can also explain the interaction between self-esteem and second language output. Self-esteem (an internal attitude) is subconscious and relates to the acquisitional side of the model. Its interaction with subconsciously acquired rules may either facilitate or block the Monitor's competence in its conscious application of the rules and thereby affect second language output.

To date, no known research has been done relating self-esteem to second language behavior. Brown (1973), suggests self-esteem as one of several areas of needed research in second language acquisition.

My focus in this study will be on the relationship between a student's self-esteem (also referred to as self-concept) and his oral production of a second language. I have chosen to focus on the oral production of a second language rather than on other skills for two reasons. First, current research by Richard Tucker on affective variables has shown that affective variables more closely relate to a student's oral production and listening comprehension than to reading and writing (presentation by Richard Tucker, University of Michigan, April, 1977). Furthermore, Gardner and Lambert (1972, p. 130) found that a relationship exists between integrative motivation (an affective variable) and second language proficiency, "especially in the oral-aural features of proficiency." This phenomenon seems natural to me since speaking is an active skill which requires risking evaluation by others of the speaker's grammar, pronunciation, language facility, and often personal worth. Second, my work with students at the English Language Institute has led me to believe that a student's evaluation of himself and his speech has an effect on his oral performance. I would like to explore my observation further.

Theoretical Framework

The purpose of the present study is (a) to focus on the relationship between self-esteem and the oral production of a second language and (b) to experiment with different measures of self-esteem.

The study of self-esteem requires the consideration of these three aspects: an adequate definition of the term; the relationships among self-esteem, academic performance, and attitudes; and measures of self-esteem.

Definitions. Rogers (1951) and Coopersmith (1967) define self-esteem in the following ways. First, Rogers (1951, p. 136):

> The self concept or self-structure may be thought of as an organized configuration of perceptions of the self which are admissible to awareness. It is composed of such elements as the perceptions of one's characteristics and abilities; the precepts and concepts of the self in relation to others and to the environment; the value qualities which are perceived as associated with experiences and objects; and goals and ideals which are perceived as having a positive or negative valence.

Furthermore, Rogers states (p. 191) "behavior is consistent with the organized hypotheses and concepts of self-structure." Coopersmith (1967, pp. 4-5) defines self-esteem in the following manner:

> By self-esteem, we refer to the evaluation which the individual makes and customarily maintains with regard to himself; it expresses an attitude of approval or disapproval, and indicates the extent to which an individual believes himself to be capable, significant, sucessful, and worthy. In short, self-esteem is a personal judgment of worthiness that is expressed in the attitudes that the individual holds towards himself. It is a subjective experience which the individual conveys to others by verbal reports and other overt expressive behavior.

Coopersmith states further that overall self-esteem is essentially resistant to change and remains fixed; unless the person actively moves to change it through therapy or some other conscious effort. However, self-esteem may vary across various areas of experience and according to age, sex, and other role defining conditions. Our sources of self-esteem are our personal experiences with ourselves, others, and our own independent judgments of the external world.

Therefore, we can say that self-esteem is an assessment or evaluation which we make of ourselves and our abilities. We express this feeling towards ourselves in ways consistent with our verbal messages and overt behavior. Our self-concept may vary according to various situations and tasks which we perform. Certainly my own self-concept is higher in situations where I am speaking English than those where I am speaking French. We can identify five external frames of reference of self-esteem (moral-ethical, physical, familial, personal and social) and three internal frames (identity, self satisfaction, and behavior), (Fitts, 1965).

Furthermore, we can also distinguish among three levels of self-esteem: global, specific, and task. Simpson and Boyle (1975) refer to global self-esteem as the individuals evaluation of his overall worth as a person. Specific self-esteem is evaluations made in certain life situations (social interaction, work, education) or those based on particular aspects of the individual (personality, intelligence, etc.) The third level is task self-esteem or the individual's expectations/evaluations of himself in task situations. The relationship of task to specific to global self-esteem is hierarchical with task as the base and global as the apex of the self-concept triangle (Shavelson et al., 1976, p. 413).

Self-esteem and its relationship to academic performance. Research on self-esteem has typically been done within the realm of psychotherapy (i.e. "before and after" testing of clients in treatment) or in terms of classifying normal and abnormal behavior. However, growing research is being done relating self-esteem to the fields of learning and education. The following are three broad areas in which such research is occurring: self-esteem and its relationship to the prediction of future performance; self-esteem and its relationship to academic performance; and self-esteem and its relationship to motivation and attitudes. Fitts (1972) reports a rather complete account of work relating self-esteem to predicted future performance, academic performance, and attitudes.

He and his colleagues have done extensive research using his self-concept measure, the Tennessee Self Concept Scale (TSCS), to predict the future performance of a wide variety of populations including child care workers, blind college students, the handicapped, school dropouts, and even paratroop trainees. His results show that self-concept, as he measures it, predicts who will continue in various training programs and who will perform well on the job once training is completed. Furthermore, work by Brodkey and Shore (1976), using an original personality test based on the Myers-Briggs style statements and using the Q-sort technique, showed that their test was strongly predictive of good and poor language study behavior as judged by language teachers. Although the Brodkey-Shore test was not defined as a self-concept test, but rather as a personality test, it utilizes a variety of personality factors, including some which are classifiable as self-concept characteristics.

In relating self-concept to academic performance, Fitts (1972) reports that work in the area of self-concept and academic performance shows that the student's concept of self is a significant variable (p. 43). However, when relating self-concept to more narrow criteria, such as achievement tests and grade point average, he has found the relationship is neither clear nor conclusive. He also states that specific self-perceptions are better predictors of grades and achievement test scores than his general measure. Simpson and Boyle (1975) found that essentially global self-esteem was unrelated to predictions of performance and to actual performance. On the other hand, specific self-esteem (in this instance defined as educational and intellectual esteem) was more closely related to and predictive of actual performance than global self-esteem and that task self-esteem (collected before and after a mid-term exam) was the most predictive of actual performance.

In relating self-esteem to affective variables, Fitts (1972, pp. 36-42) reports research that shows a relationship exists between self-concept and factors affecting academic performance (attitudes toward school and teachers, assuming responsibility for learning, motivation and goals, morale and satisfaction with school, class participation, discipline problems, dropout rates, school failures, etc.) He concludes that a person with a healthy concept (1972, p. 43) is "apt to use his intellectual resources more efficiently and this may be a critical factor in his achievement if his intellectual resources or educational background are borderline. Otherwise his self concept seems to be more closely related to the noncognitive aspects within the academic setting."

Measures of self-esteem. There are three major techniques used to measure self-esteem: direct self-evaluation; self-ideal discrepancy scoring; disguised self-evaluation. Briefly, direct self-evaluation asks the subject to either agree, disagree, or remain neutral on statements which are favorable/unfavorable to himself. The self-ideal discrepancy asks the subject to rate himself as he feels he is and as he would like to be. Self-esteem is inferred from the similarity of the two descriptions. The third type is a disguised self-evaluation which requires the subject to evaluate an aspect of himself (i.e. handwriting, picture, voice) which is distorted so that he does not consciously recognize the source of the stimulus. Self-esteem is measured by the similarity of ratings between his self rating and the distorted rating. A more descriptive review of the theoretical assumptions, characteristics, validity, and analyses of self-concept measures appears in the following sources: Wylie (1961); Taylor and Reitz (1968); and Robinson and Shaver (1973).

Definitions and Hypotheses

For the purposes of this study, I have chosen the following definitions (adaptations from Rogers, 1951; Coopersmith, 1967; and Simpson and Boyle, 1975) and hypotheses.

Global self-esteem refers to those evaluations of our worthiness which we consciously make of ourselves. Such evaluations are consistent with our belief system, our verbal behavior, and our nonverbal behavior. Facets of global self-esteem which will be evaluated are the moral-ethical self, the personal self, the familial self, the physical self, the social self, identity, self-satisfaction, and behavior. Subjects of high global self-esteem will be operationally defined as those scoring one standard deviation above the mean or above the 82 percentile on the total positive scale of the Tennessee Self Concept Scale. Subjects of low self-esteem will be operationally defined as those scoring more than one standard deviation below the mean or below the 18th percentile. Furthermore, each group of subjects will be cross checked for high self-criticism scores which indicate inflated total self esteem scores. Their scores will be deleted from the population.

Specific self-esteem refers to evaluations subjects make of themselves in situations where they are using English and to evaluations they make of individual aspects (i.e. language learning ability, intellectual ability).

Task self-esteem refers to evaluations subjects make of their performance before and after an extemporaneous speech in English or in French/Spanish.

My hypotheses are the following: 1. Global self-esteem (GSE) scores will correlate positively with specific self-esteem (SpSE) scores and specific self-esteem scores will correlate positively with task self esteem (TSE) scores. 2. TSE scores will be most closely correlated with performance on an extemporaneous speech; SpSE scores will be less closely correlated with performance on an extemporaneous speech; and GSE scores will have the lowest correlation. 3. High self-esteem subjects (HSES) will receive significantly higher scores from their teachers on their oral performance than moderate esteem subjects (MSES) and MSES will score significantly higher than the low self-esteem subjects (LSES). I also anticipate that HSES will rate themselves higher on their oral performance than MSES and MSES will rate themselves higher than LSES. 4. SpSE measures will have a stronger relationship than GSE measures to predicted performance on the extemporaneous speech.

Although research in the area of self-esteem and language acquisition appears to be non-existent, similar research in a related field (self-esteem and performance in a psychology class) by Simpson and Boyle (1975) showed a relationship between the three levels of self-esteem and performance on a mid-term exam. Using students enrolled in a college psychology course, they showed (a) global self-esteem related positively to specific self-esteem and that specific self-esteem related positively to task self-esteem; (b) that task self-esteem was the most closely related to performance on a mid-term exam, specific self-esteem was less closely related, and global the least closely related; (c) that high esteem subjects received significantly higher scores than mid esteem subjects and mid esteem subjects had higher scores than low esteem subjects; and (d) that specific self-esteem scores had a stronger relationship than global self-esteem scores to predicted performance.

Pilot Study

During the spring of 1977, I conducted a pilot study exploring the relationship between global and specific self-esteem and the oral production of English as a Second Language. Subjects were students enrolled in levels 150 and 160 of the intensive course at the English Language Institute. Hypotheses included the following: (1) global self-esteem (GSE) would correlate positively to specific self-esteem (SpSE); SpSE would be more closely related to teacher and self ratings of oral productionthan GSE ratings; (3) High self-esteem subjects (HSES) would receive higher teacher oral production ratings than the mid self esteem subjects (MSES) and the MSES would receive higher scores than the low self esteem subjects (LSES). In addition, HSES would rate their self oral production higher than MSES and MSES would rate themselves higher than LSES; (4) total self-esteem scores would be predictive of the "good" language learners of my population as well as the "bad" learners as rated by teachers and themselves on their oral production.

The global self-esteem measures consisted of a single-item self-esteem scale and an adaptation of the Sherwood Self Concept Inventory (Robinson and Shaver, 1973). The specific self-esteem measure consisted of a single item which asked subjects to rate their language learning ability on an eight point scale that used good and poor language learning ability as anchor points. Oral production ratings were obtained through teacher and self ratings.

Results showed some indication that the two levels of self-esteem, global and specific, appeared to be somewhat related and that there may be a tendency for specific self-esteem to be more closely related to oral performance than global self-esteem. In addition, high self-esteem subjects received higher oral production ratings from themselves and their teachers than the low self-esteem subjects. Finally, the results showed a tendency for specific self-esteem to be more closely predictive of oral performance than global self-esteem. The results of this study are inconclusive since my population was small (N = 15) and data was collected for only two levels of self esteem (global and specific.) This research shows the need for further investigation in the area of self-esteem and the oral production of a second language.

Proposed Research

My proposed research deviates from that of Simpson and Boyle in the following ways. First, the two studies have different populations. Theirs focused on college students studying psychology while mine focuses on adults (college age and older) learning a second language. Second, their study deals with two types of specific self-esteem (intellectual and educational) while mine focuses on a wider variety of self-esteem areas.

Figure 1 presents a model of the Simpson and Boyle study.

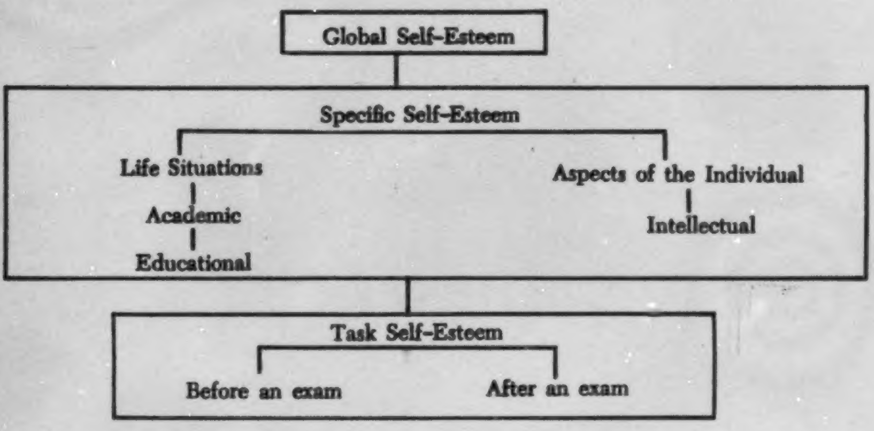


Figure 1. Levels of self esteem as defined by Simpson and Boyle (1975).

Through the use of the subscores of my global self-esteem measure and the addition of measures particular to language learning, I am expanding this model to include a wider variety of specific self-esteem factors (see Figure 2). Global self-esteem is a combination of several kinds of specific self-esteem in my model. The general topic of specific self-esteem consists of evaluations made in life situations and those based on certain aspects of the individual. For the purposes of this study, life situations will be concerned with only the two areas of community interactional and academic related situations. Third, Simpson and Boyle divide their population in to thirds in order to obtain equal numbered high, mid, and low self esteem groups while my three groups will be divided according to their distance from the mean. Fourth, Simpson and Boyle do not use the self criticism scores and defensive positive scores of the Tennessee Self Concept Scale in their study while I intend to include them in my study. The self criticism score serves as a cross check on total self-esteem scores by indicating which scores of the population are inflated due to defensiveness and a need to look better on paper for the researcher. The defensive positive score serves the same purpose on a more subtle level. Fifth, my study

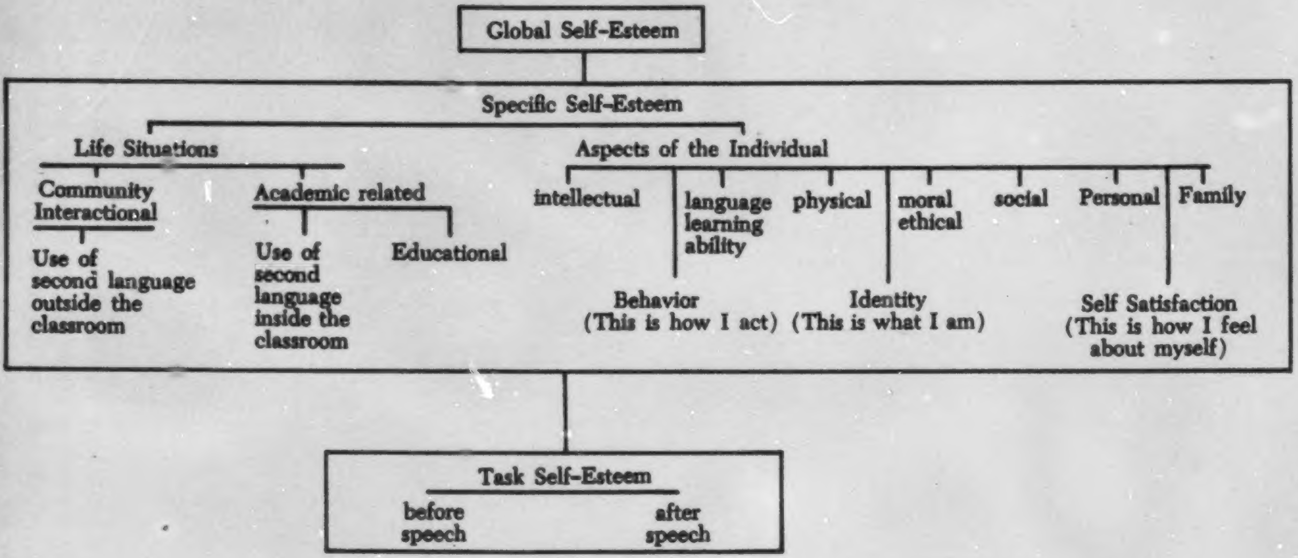

Global Self-Esteem
Specific Self-Esteem
Life Situations
Community Interactional
Use of second language outside the classroom
Academic related
Use of second language inside the classroom
Educational
Aspects of the Individual
intellectual
language learning ability
physical
moral ethical
social
Personal
Family
Behavior (This is how I act)
Identity (This is what I am)
Self Satisfaction (This is how I feel about myself)
Task Self-Esteem
before speech
after speech

Figure 2. Levels of self esteem examined in the present study.

contains a population from a variety of cultures learning one of two languages: English or French/Spanish (one language to be specified before the study begins).

Subjects

Subjects will include two groups of students. Group I will consist of students enrolled in courses in English as a Second Language during Term I, 1977, at the English Language Institute. The tests will be administered to students enrolled in levels 140, 150, and 160. In order to complete the Tennessee Self Concept scale, subjects must have at least a sixth grade reading level. I intend to administer the Stanford Achievement Test for Reading, Intermediate 2 in order to determine which students do not have at least a sixth grade reading level. Those subjects who have not attained the appropriate reading level will not be included in the study. Additional samples of ESL learners may be drawn from similar institutes. Group II of my population will consist of students enrolled in the beginning level (101) of French or Spanish at the University of Michigan during Fall of 1977.

Self-Esteem Measures

Global Self-Esteem. I intend to use the Tennessee Self Concept Scale (Fitts) to measure global self-esteem. This scale consists of 100 statements which the subject marks on a five point scale ranging from completely false (0) to completely true (5). There are three internal frames of reference: identity, self-satisfaction, and behavior which relate to these five aspects of self: moral-ethical, physical, social, familial, and personal. In addition, self-criticism scores will be computed to weed out subjects who have inflated their self-esteem scores.

The test-retest reliability for the various subscores of the Tennessee Self Concept Scale falls in the range of .70 to .90 for American adults, age 12 and up (Robinson and Shaver, 1973). The test-retest reliability for the total positive self-esteem scores is .92 for the same population. Evidence so far shows no need to establish separate norms for age, sex, race, or other variables (Fitts, Manual, 1965, p. 13). Furthermore, Fitts (1971) reports the norms for other cultures seem comparable to those for Americans. Unfortunately, he does not report the reliability scores for other cultures.

Specific Self-Esteem. To measure specific self-esteem, I intend to use five statements based on the following: two life situation statements pertaining to English use, one related to language learning ability, and one question each related to educational and intellectual esteem taken from the Simpson and Boyle study. The two questions related to intellectual and educational self-esteem had a test-retest reliability of .84 and .81 respectively over a two month period of time. Copies of the questionnaire are attached. The other specific self-esteem items (physical, moral-ethical, familial, personal, social, self-satisfaction, behavior, and identity) are subscores of the Tennessee Self Concept Scale with test-retest reliabilities over a two week period of time of .87, .80, .89, .85, .90, .88, .88, and .91 respectively (N = 60 college students).

Task self-esteem. I intend to use two measures of task self-esteem. Measure one will ask students to estimate the judges' score of their oral production before they have given an extemporaneous speech. Measure two will ask them to estimate the judges' score of their oral production after they have given the speech. Both measures will use a seven point scale ranging from 0 (not nativelike) to 7 (near nativelike). Copies of the questionnaire are attached.

Oral Production Task and Ratings

Oral Production Task. The oral production task will require students to give a two minute extemporaneous speech on his choice of one of two topics: "Three places I would take a Foreigner to visit in my country" and "My plans for ten years from now." Each subject will have approximately three minutes to prepare his speech before presentation. The presentation will be recorded on cassettes. Sheets for the task are attached.

Oral Production Ratings. I intend to collect three different ratings of student oral production of English and French/Spanish: a student self rating; a combined teachers' rating; and a rating of a speech sample by three judges. The self rating scale will ask the student to rate his oral production on a seven point scale ranging from near nativelike (7) to non nativelike (0). Samples of this and the other rating scales are attached. The teacher rating scale will ask the teachers to rate their students' oral production on a seven point scale ranging from 7 (near nativelike) to 0 (non nativelike). Teacher scores will be averaged in order to obtain the teacher oral production score. Three judges, all native speakers of English or French/Spanish depending on the population being judged, will be asked to rate a two minute segment of each student's extemporaneous speech. The judges' rating scale will be the same as the self and teacher rating scales. Their scores will also be averaged.

Analysis of Results

I plan to use the following statistical measures.

1. The multiple regression technique will account for the different levels of learning among subjects in my sample. Learning refers to the knowledge students have of the language. Those scores falling above the regression line will indicate high self-esteem and high oral production students and those falling below, the low self-esteem and low oral production students.

2. Independent "t" tests and product moment correlations will be run to reveal what, if any, sex and cultural differences occur.

3. A factor analysis of the correlation matrix using the principal components method with a varimax rotation will extract those factors accounting for the variance.

4. Multiple product moment correlations will reveal that the patterns of

relations among the various levels of self-esteem (both between levels and within levels) are as predicted.

5. A one-way analysis of variance performed on the oral production task scores, the teacher scores, and the self oral ratings of the three esteem groups (high, mid, low) should reveal significant effects of the various measures of self-esteem.

6. A one-way analysis of variance performed on the predicted scores of the high, moderate, and low esteem groups on the oral performance task should reveal significant effects for at least the specific self-esteem groups.

Implications of the Results and Limitations of the Study

I anticipate the hypotheses to be supported and the result to have application to futher research in the area of self-esteem and language acquisition.

Two limitations of studies involving affective variables are the tests used and the discrepancy between the language of the test and the language of the students. This study has both limitations. I have tried to choose the best available measures for self-esteem. The Tennessee Self Concept Scale, which accounts for the global measure and several areas of the specific measure, has been used in numerous studies (Fitts, 1965) and appears to be the best established self-esteem test (Robinson and Shaver, 1973). The intellectual and educational specific measures have respectable test-retest reliabilities. The new specific items (language use inside and outside the classroom and language learning ability) will have test-retest reliabilities established prior to the collection of data.

The second problem of this type of study is whether or not to translate the tests into the native language of the subjects. Translations of the Tennessee Self Concept Scale exist and norms have been established for some of the language groups to be used in this study; however, as in any written text, there are problems of translation equivalence. Fitts (personal communication) points out the problems of translating the true meaning of idiomatic expressions as well as the problems of translating the appropriate cultural meanings and the social desirability of items. Furthermore, once the test has been translated, measurement problems of norms, reliability, and validity must be considered. To date only the Hebrew version of the TSCS has been thoroughly tested enough to provide valid and reliable results. I have chosen, therefore, to use the English version of the scale. By limiting my population to subjects who are in the higher levels (140, 150, and 160), I have tried to choose a population that can read and understand the scale without the need for translation. Moreover, a portion of the population (the subjects studying French/Spanish) will be adult native speakers of English and will not have problems with the language of the test.

The results of the study will relate to other research being done with affective variables at the University of Michigan and will contribute to a more composite picture of the good or poor language learner. Additional research might examine the relationship between self-esteem and motivation and to relate self-esteem to the learning of other second languages. Eventually, I hope this research will be applicable to the development of methods of increasing students' self-esteem, especially at the specific task levels, and thereby improve their oral production of a second language.

REFERENCES

Abel, Theodora M. 1973. *Psychological testing in cultural contexts.* New Haven, Connecticut, College and University Press.

Barksdale, L. S. 1972. *Building self-esteem.* Idyllwild, California, Barksdale Foundation for Furtherance of Human Understanding.

Bills, Robert E., E. L. Vance, and Orison S. McLean. 1951. An index of adjustment and values. *Journal of Consulting Psychology 15,* 257–261.

Brodkey, Dean, and Howard I. Shore. 1976. Student personality and success in an English language program. *Language Learning 26,* 1, 153–159.

Brown, H. D. 1973. Affective variables in second language acquisition. *Language Learning* 23, 2, 231–244.

Chun, K-T, Sidney Cobb, and John R. P. French. 1975. *Measures for psychological assessment.* Ann Arbor, Michigan, Institute for Social Research.

Clement, R., Louise H. Major, R. C. Gardner, and P. C. Smythe. 1977. Attitudes and motivation in second language acquisition: an investigation of Ontario francophones. *Working Papers in Bilingualism 12,* 1–20.

Coopersmith, Stanley. 1967. *The antecedents of self-esteem.* San Francisco, W. H. Freeman and Co.

Curran, Charles. 1972. *Counseling learning: a whole person model for education.* New York, Grune and Stratton.

Evans, Richard I. 1975. *Carl Rogers: The man and his ideas.* New York, E. P. Dutton and Co., Inc.

Fitts, William H. 1972. *The self concept and performance.* Nashville, Tennessee, Dede Wallace Center.

———. 1971. *The self concept and self-actualization.* Nashville, Tennessee, Counselor Recordings and Tests.

———. 1965. *Tennessee self concept scale manual.* Nashville, Tennessee, Counselor Recordings and Tests.

Jersild, Arthur T. 1952. *In search of self, an exploration of the role of the school in promoting self-understanding.* New York, Bureau of Publications, Teachers College, Columbia University.

Miskimins, R. W., and George Braucht. 1973. *Description of the self.* Colorado, Rocky Mountain Behavioral Science Institute.

Naiman, N., M. Frohlich, and H. H. Stern. 1975. *The good language learner.* Toronto, Canada, Ontario Institute for Studies in Education.

Robinson, John R., and Phillip R. Shaver. 1973. *Measures of social psychological attitudes.* Ann Arbor, Michigan, Institute for Social Research, University of Michigan.

Rogers, Carl. 1970. *On encounter groups.* New York, Harper and Row.

———. 1961. *On becoming a person.* Boston, Houghton Mifflin.

———. 1951. *Client centered therapy.* Boston, Houghton Mifflin.

Shavelson, Richard J., Judith J. Hubner, and George C. Stanton. 1976. Self-concept: validation of construct interpretations. *Review of Educational Research 46,* 3, 407–441.

Simpson, Kenneth, and David Boyle. 1975. Esteem construct generality and academic performance. *Educational and Psychological Measurement 35,* 4, 897–904.

Taylor, James B., and Willard E. Reitz. 1968. *Three faces of self-esteem.* Research Bulletin 80. London, Canada, University of Western Ontario, Department of Psychology.

Wylie, Ruth. 1961. *The self concept.* Lincoln, Nebraska, University of Nebraska Press.

APPENDIX A

SPECIFIC SELF-ESTEEM QUESTIONNAIRE

Student Questionnaire

Name______________________________ Level______________ Section______________

Language__________________________ Country__________________________

Instructions: Place an "X" on the line which correctly answers the question.

1. Generally, how do you rate your ability to learn a language?

_____ _____ _____ _____ _____ _____ _____ _____

0 7

Extremely low Extremely high

2. Generally, how do you feel about yourself when you speak English in the classroom?

_____ _____ _____ _____ _____ _____ _____ _____

0 7

Very bad Very good

3. Generally, how do you feel about yourself when you speak English outside the classroom (i.e. in the dorm, at the store, with American friends)?

_____ _____ _____ _____ _____ _____ _____ _____

0 7

Very bad Very good

4. Generally, how do you feel about your intellectual abilities?

_____ _____ _____ _____ _____ _____ _____ _____

0 7

Very bad Very good

5. Generally, how do you feel about your educational-academic abilities (in classes and other situations directly related to your education)?

_____ _____ _____ _____ _____ _____ _____ _____

0 7

Very bad Very good

6. How do you rate your overall proficiency of English?

_____ _____ _____ _____ _____ _____ _____ _____

0 7

Not nativelike Near nativelike

7. How do you rate your speaking ability of English?

_____ _____ _____ _____ _____ _____ _____ _____

0 7

Not nativelike Near nativelike

APPENDIX B

Student Speech

You are to choose 1 of the 2 topics listed below and give a short 2 minute speech on it. You will have 3 minutes to prepare your speech. You may make notes in the blank space below.

My plans for ten years from now
Three places I would take a foreigner to visit in my country

APPENDIX C

Student Task Questionnaire

Name________________________________ Section______________
Language______________________________ Level______________
Country_____________________

What score do you feel the judges will give you on your speech?

____ ____ ____ ____ ____ ____ ____ ____

0 Not nativelike	7 Near nativelike

NOTE: The same form will be used before and after the speech is given.

APPENDIX D

Judges' Rating Sheet

Instructions: Listen to each short segment of speech and then mark the student's level of oral production on the scale at the right of his name. A score of 0 means that you feel the student's oral language is non nativelike and a score of 7 means that you feel the student's language is near nativelike.

Student's Name	Oral Production Rating
	____ ____ ____ ____ ____ ____ ____ ____
	0 Not nativelike 7 Near nativelike

APPENDIX E

Teacher Questionnaire

Instructions: Please list your students alphabetically in the column on the left. Rate their oral production (speaking ability) in the column on the right. Please return to Adelaide Heyde as soon as possible.

Student's Name	Oral Production Rating
	____ ____ ____ ____ ____ ____ ____ ____
	0 Not nativelike 7 Near nativelike

Diary of a Language Learner: An Introspective Study of Second Language Learning

Francine M. Schumann and John H. Schumann

University of California, Los Angeles

It is now well-known that success in second language learning (SLL), particularly in the environment where the target language is spoken, depends to a large extent on the learner's attitude, motivation and cultural adjustment. Especially important is the resolution of such inevitable problems as language shock and culture shock. In the past these social-psychological aspects of SLL have been examined in cross-sectional studies in which language learning by large groups has been studied. What has been lacking have been in-depth longitudinal case studies examining the social-psychological variables in SLL by *individuals*.

This paper reports on the first stage in the development of a methodology for conducting such longitudinal case studies. Two adult second language learners kept intensive journals during three identical language learning experiences: as beginners attempting to acquire Arabic in North Africa, as intermediate learners studying Persian at UCLA and again as intermediate learners acquiring Persian in Iran. In the journals the learners recorded their feelings and reactions toward the foreign cultures, the target language speakers and the methods of instruction. The detailed records of feelings and reactions contained in the journals indicate that for each individual there are *personal* variables that can either promote or inhibit SLL. In this paper such personal variables are identified (i.g. nesting patterns, transition anxiety, rejection of teaching methodology, maintaining a personal agenda in language learning). In addition, the authors discuss the use of journals not only as a technique for examining the social psychological aspects of SLL but also as a possible vehicle for facilitating the language learning process.

This paper addresses itself to the study of individual language learning experiences in order to see what factors affect the individual's progress. The authors are both the subjects and the researchers in this study. The raw data for the study was collected in the form of a diary. Below, the authors attempt to present their rationale for this type of study followed by a step by step account of the procedures used in writing the diaries and the results thus far obtained from initial investigation of the data. These results are in no way complete; therefore, this is a report of work in progress. Further analysis will hopefully provide additional insights into an individual language learner's learning strategies.

It is slowly being recognized that success in second language learning (SLL), particularly in the environment where the target language (TL) is spoken, depends to a large extent on a variety of social-psychological variables. In the past these variables have been examined in cross-sectional studies, often using factor-analytic techniques, in which language learning by large groups has been studied. In these studies only a limited number of variables have been examined—those which can be measured and those for which measurement instruments exist. Therefore, what is said about SLL of a particular group is only true of a *group* of individuals. Within that group each individual has his own social-psychological profile of learning. What has been lacking in SLL research have been in-depth longitudinal case studies examining the social-psychological profile in SLL of an *individual*. This project is an attempt to provide such research. It addresses itself to the study of individual language learning experiences to see how various social-psychological factors affect an individual's perception of his own progress.

If one takes a global view of epistemology, one can see that traditional psychologies of the East (such as zen, yoga and sufism) examine a phenomenon by having individuals experience that phenomenon. One learns things by going through an experience and by observing oneself during that experience. In Western psychology, particularly in experimental psychology, the experimenter rarely observes himself. He observes others undergoing a particular process and uses various kinds of data-gathering procedures during this period of observation. Breaking away from the methods of experimental psychology, the authors have focussed on self-observation as a research tool and a source of scientific data.

The opportunity for a self-observational study of SLL presented itself when the authors spent two months during the summer of 1976 in Tunisia and another two months during January and February of 1977 in Iran. Also included in the study is their participation in an intermediate Persian class at UCLA during the fall of 1976. In Tunisia they were learning Arabic at the beginning level. John studied Tunisian Arabic in a class which met one hour a day, and Francine studied Modern Standard Arabic in a class which met four hours a day. The instructional program lasted six weeks. In Iran they pursued the study of Persian at the intermediate level, working on their own without formal instruction. Both of them had previously spent three years in Iran in the Peace Corps.

In order to document what actually was involved in their attempts at learning these second languages, they decided to experiment with the use of diaries as research instruments. The plan was to keep detailed journals of all cross-cultural and linguistic events which each of them perceived as relevant to their SLL. To prepare themselves for this task they attended a journal workshop at UCLA during May of 1976. The format for this workshop was developed by Ira Progoff (1975) and is designed to train the participant to examine his life experience from approximately twenty different perspectives and then to integrate them into a unified whole. The workshop was useful in that it provided the authors with a sense for the various ways a journal can be used to document and understand an experience. However, in the actual research they did not attempt to rigorously follow the Progoff format. Instead, they found it more suitable to record daily events and the thoughts and feelings related to them in a log-like fashion, paying particular attention to cross-cultural adjustments and efforts made and avoided in learning the TL both in and out of class. Because the diaries must be as candid as possible and entries are uncensored, they are essentially private documents. Therefore, for the study this raw data was rewritten by each subject keeping all relevant detail but eliminating highly personal entries. Thus, the data for the project is reported in the following form:

1) The story of Francine's language learning experience in Tunisia, at UCLA, and in Iran, and
2) The story of John's language learning experience in these three places.

Each of these stories is preceded by an account of the history of the authors' involvement with second languages. It was felt the value in writing these historical accounts would be in providing a more total picture of the authors as learners by showing what languages they had studied or come in contact with during their childhood, adolescence and young adulthood.

A phenomenological analysis of the data was used to determine what social and psychological variables enhance or inhibit SLL. In conducting this phenomenological analysis, the first question asked of the data was what it revealed about the social-psychology of SLL that was not known before. This approach was not very productive. There was nothing substantial in the data that was totally new. The question was then reformulated to ask, what the data told the subjects as *individuals* about their language learning. Each of them then went through their own data and tried to identify the important variables in their language learning. It was felt this approach was very productive because it revealed idiosyncratic patterns of behavior that seriously affected their language learning and that had not been specifically identified as important in previous studies of the social-psychology of SLL.

In research such as this which is both clinical and phenomenological, results cannot be displayed in a series of figures and tables. Nevertheless, the authors have attempted to sketch out some of the things they have learned about themselves. However, due to constraints of space only a few of the authors' observations can be presented here. A detailed report of this research including both the data and analysis will be available in F. M. Schumann (forthcoming). Below Francine describes three of her behavior patterns followed by John's description of three of his.

Obsession With Nesting

In both the Tunisian and Iran journals, it is clear that in order for me to be able to devote the time, energy or emotional involvement required in language learning, I must first feel content in the place I am living. My surroundings must be orderly, comfortable and have my imprint on them identifying them as my home away from home. I devote days to putting my nest in order leaving very little energy or even thought for anything else. If I am frustrated in my attempt to build a suitable nest, it will have a negative effect on my pursuing any creative endeavor such as language learning.

In Tunisia, I was totally unable to fulfill this nesting instinct. We lived in a 300 year old building in the medina (the Arab part of town near the souk). I was never able to adjust to the lack of hot water, comfortable furniture and a suitable place to put away my belongings. I was also frustrated by the constant layers of dirt blowing in and having to put up with a strange type of bug in the house that attacked us all summer causing us to look like smallpox victims throughout our stay. The result was that I felt alienated from and hostile towards my environment, and these emotions usurped my energies such that I rarely could direct any to studying Arabic the entire time we lived there.

Again this nesting instinct was an issue with me upon our arrival in Iran. Fortunately, the accommodations there lent themselves to my fulfilling this instinct and doing so within a reasonable period of time. Because we lived in a hotel, I was free of cleaning responsibilities and the room was put in order for me daily. The furniture both accommodated our belongings and was comfortable and even provided a very suitable work area. Thus, it was only a matter of unpacking and deciding on convenient places to put things away. Within a week of our arrival I was content that things were to my satisfaction. I was then able to begin to direct my energy to the task of studying Persian.

Withdrawal From Learning Due to Nonacceptance of the Teaching Method

A rigid adherence to the audio-lingual method was followed in my Arabic class. The TL was used exclusively as the medium of instruction. Weekly lessons included full page dialogues followed by approximately 6-8 pages of drills. No translations, grammatical explanations or vocabulary lists were given during the lessons nor did the text provide one with these supports. In class students were forbidden to write anything down to aid them in remembering, to copy words or sentences off the chalkboard, to look up things in the text or to consult with fellow learners for clarification. The rule was listen, repeat and respond over and over for four hours. The sole source of the language was the teacher. I hated the method. My anger bred frustration, a frustration which I acutely felt as my goal was to be a star performer in class, and I found it impossible to be so under these circumstances.

Instead of resorting to a solution which would allow me to cope with this learning environment and learn in spite of the method, my reaction was to reject it and withdraw from learning. This withdrawal was gradual and displayed itself in a variety of ways. Some days I would assume such a low profile in class, making no attempts to participate, that only my physical presence allowed that I was indeed a member of the group. Other days this withdrawal took the form of my cutting-up during the lesson. Eventually the withdrawal led to my leaving class early, walking out on exams and on some days not showing up at all.

Motivation for Choice of Language Materials

The subject matter and packaging of language materials have a strong effect on whether I will make use of them as learning resources. I had decided that in order to advance in the acquisition of Persian it was necessary for me to learn to read the language. A wide selection of reading materials for students of Persian does not exist, especially if one is working on a program based on self-instruction. The materials that were available to us for this task were the elementary and intermediate *Modern Persian Readers* published by the University of Michigan. These materials contained reading selections mainly from newspapers. In the elementary book, the text is presented in Persian script on one page and in phonemic transcription on the facing page. Each selection is accompanied by an extensive vocabulary list. The book is really quite suitable for self-study by someone who has a basic knowledge of spoken Persian. However, I found I could not study from these books because 1) I found the subject matter boring (for example, the arrival of the West German minister of agriculture at the Tehran airport, what government officials met his plane and what was said to welcome him), and 2) because the book itself was unattractive.

In Tehran they sold books for children in Persian that had fantastic art work. It was this art work which greatly attracted me, and I decided I wanted to learn to read using these children's books. John said I shouldn't waste money on them as they contained no vocabulary lists or transcriptions (which he liked in the Michigan readers) and therefore, all my time would be spent looking up words in the dictionary and mispronouncing a good part of what I read. Finally, at the end of the first month of our stay John agreed to try some.

To my delight we had enormous success with them. The stories were so entertaining that they led to a much greater effort learning to read than did the Michigan readers. In fact, studying now became that part of the day that we looked forward to as a special treat. The illustrations greatly aided our understanding of much of the action. We underlined words which we were unsure of and later consulted an Iranian friend about their meanings. An added bonus was the insight these stories gave us to Iranian culture (for example, the bond of loyalty and friendship among children, the stories that are revealed in the designs on Persian carpets, and the awe and respect the Iranian child has for the grandfather.) In sum, had I not been so drawn initially by the beautiful art work we might never have experienced the success or pleasure of learning to read that these children's books ultimately provided.

The following are factors that John perceived as affecting his language learning:

Transition Anxiety

The Tunisian and Iran experiences taught me that I am most vulnerable to culture shock during transitions: i.e. entering the foreign country, leaving it and changing residences while in that country. I find that the details of these transitions (making reservations, buying tickets, getting to the airport, going through customs, finding a hotel, moving to and setting-up an apartment) are emotionally draining and anxiety provoking. Once I am there, I am quite content (even in the absence of many creature comforts), but when I begin to anticipate the next transition, the emotional stress starts again.

I found one reasonably effective way to control this stress during travel to the foreign country. Enroute to Tunisia and during the first week or so after arrival, I devoted every free minute to working my way through an elementary Arabic reader. On the way to Iran I kept myself occupied by working through a low-intermediate Persian reader. I believe these were productive solutions because they allowed me to learn more of the language while dealing with the anxiety. Learning more of the language gave me a sense of satisfaction and accomplishment that went a long way toward counter-acting the anxiety. This approach, however, did not work when leaving the foreign country since during departures, I did not feel motivated to study because I wouldn't be needing the language back in the United States.

The Desire to Maintain a Personal Agenda in Language Learning

I discovered that I like to have my own agenda in second language learning. In other words, I like to do it my way. However, I also found that my agenda is often in conflict with the teacher's. Since the teacher's agenda is the standard for the class, then my not following it is failure in the eyes of the other class members. In order not to look like a failure, I have a tendency to compromise my agenda in the direction of the teacher's. For example, in Tunisia, the teacher wanted the students to memorize the dialogues perfectly. I felt that to learn them that well was unnecessary and that it required spending time that could be put to better use in other language learning activities. But, If I did it my way and did not memorize the dialogue as perfectly as possible then I would be frustrated and embarrassed because my poor performance would be judged by the class according to what the teacher wanted the students to accomplish and not what I wanted to accomplish.

My desire to pursue my own agenda was somewhat more successful in the Persian class at UCLA. There the teacher occasionally asked the students to write, but I was only interested in speaking and reading, not writing. In this class I could choose not to do a writing assignment both because I was an auditor and because the teacher was flexible enough to allow it. If on occasion I were forced to write (e.g. during an in-class vocabulary quiz), my failure at this task would not be public, it would simply be between the teacher and myself. Therefore, I would not be embarrassed and could stick to my own language learning agenda.

As a result of my desire to learn a language my own way, I find I like low key teachers who don't demand too much. What I want in a teacher is someone whom I am confident in, who is an able guide and resource person, but who is flexible enough to allow reasonable latitude for me to achieve my own language learning goals.

Eavesdropping vs. Speaking as a Language Learning Strategy

As a general learning strategy, I prefer eavesdropping to speaking as a way of getting into a language. Some people learn languages by being very socially out-going and by talking with TL speakers. I prefer to get input by listening to the TL without having to become involved in a conversation. This tendency manifested itself in several ways. In Iran I enjoyed listening to the radio, watching television and participating in events where I could hear people speak without having to speak myself. Thus, at faculty meetings, at certain social gatherings and while observing classes that were conducted in Farsi, I listened as attentively as possible and wrote down words which I did not understand. Later I would learn the meanings of these words by looking them up in a dictionary or asking a friend. I do not think this strategy excludes having an integrative motivation. In my case, it was a natural outcome of being integratively motivated, but shy. Unfortunately, most language teaching methods do not allow this approach. I was able to employ it in Iran where I was learning on my own, but I could not use it in Tunisia where I was taking classes.

In conclusion, having asked what the diary study taught them about their own language learning, the authors turned once again to the original question: "What has the journal study revealed about SLL that wasn't known before?" They now think the answer is that the journal has revealed that there are numerous PERSONAL variables (a term suggested by Amado Padilla) that affect the acquisition of a second language. These personal variables, to name but a few, are issues such as those described above:

1. nesting patterns
2. reactions to dissatisfaction with teaching methods
3. motivation for choice of materials
4. transition anxiety
5. desire to maintain one's own language learning agenda, and
6. eavesdropping vs. speaking as a language learning strategy

Previous research identified and investigated psychological variables such as attitude, motivation, language shock and culture shock (Gardner and Lambert, 1972; Schumann, 1975); social variables involving dominance patterns and integration strategies (Schumann, 1976); cognitive variables relating to cognitive development (Rosansky, 1975; Krashen, 1975) and cognitive style (Naiman, Frohlich and Stern, 1975); and personality variables such as tolerance for ambiguity, sensitivity to rejection, introversion and extroversion (Naiman, Frohlich and Stern, 1975). What the authors believe this research reveals is that there are PERSONAL variables that interact with all of the above in patterns that are idiosyncratic for each individual.

To further explicate this point of view the authors offer the following analogy. They suggest that language learning be seen as analogous to the operations of a pinball machine. The machine is the total language learning situation. The knobs represent the social, psychological, cognitive, and personality variables mentioned above. In addition, there are knobs that represent age, aptitude and instructional variables. The balls represent the learners and embody the sorts of personal variables that have been described in this paper. As each ball travels through the machine, it hits certain knobs and not others. In addition, it hits certain knobs more than others. The personal variables in each ball determine the path it takes through the machine and the extent to which it is influenced by each knob. No ball will exactly duplicate the path of any other ball, in the same way that no two learners will acquire a second language in precisely the same way.

The authors would also like to suggest that the journal research may have some practical implications. Our profession spends a good deal of time in teacher preparation, teacher training and teacher education. Perhaps second language learning might be improved by investing some time in *learner* preparation, *learner* training and *learner* education. It is possible that people such as college foreign language majors, Peace Corps volunteers and Foreign Service officers could use journals to learn about the personal variables that promote or inhibit their language learning. The goal would be for the learners to use the journal information to minimize the inhibiting aspects of these variables. The learner education program might include a course in second language acquisition which, in non-technical language, would inform the learner of the nature of his task. The goal of the course would be to dispel myths about SLL which the learner may hold and which may be counter-productive.

Finally, such a learner education program could also have a research component in which the learners would be tested on as many other relevant variables as possible, such as personality, cognitive style, language learning aptitude, hearing acuity, cerebral dominance and ego-permeability. The goal of the research would be to determine the effect of these factors on SLL and then to find ways to ameliorate them to enhance learning.

The authors add one caution however. As the profession comes to recognize more and more the importance of affective factors in SLL, there is a danger that language teachers will assume the roles of pseudo-psychologists and language classes will become group therapy sessions. While an awareness of the affective nature of SLL is essential, we must be careful not to neglect the importance of grammar, drill and the other traditional components of language instruction. In other words, as we adopt new perspectives on language learning and teaching, we have to be careful not to become excessive in any area and risk throwing the baby out with the bath water.

REFERENCES

Farzan, M., M. A. Jazayery, and H. H. Paper (eds.). 1962. *Modern Persian Reader I. Elementary.* Ann Arbor, The University of Michigan Press.

Gardner, R. C., and W. E. Lambert. 1972. *Attitudes and motivation in second-language learning.* Rowley, Mass., Newbury House Publishers.

Jazayery, M. A., and H. H. Paper (eds.). 1962. *Modern Persian Reader II. Intermediate.* Ann Arbor, The University of Michigan Press.

Krashen, S. 1975. The critical period for language acquisition and its possible bases. In D. R. Aaronson, and R. W. Rieber (eds.), *Developmental psycholinguistics and communication disorders*. New York, The New York Academy of Sciences.

Naiman, N., M. Frohlich, and H. A. Stern. 1975. *The good language learner*. Toronto, Ontario, Modern Language Center, Ontario Institute for Studies in Education.

Progoff, I. 1975. *At a journal workshop*. New York, Dialogue House Library.

Rosansky, E. J. 1975. The critical period for the acquisition of language: some cognitive developmental considerations. *Working Papers in Bilingualism* 6, 93–100.

Schumann, F. M. (forthcoming). *An introspective study of second language learning.* Masters Thesis. English Department (TESL), University of California at Los Angeles.

Schumann, J. H. 1975. Affective factors and the problem of age in second language acquisition. *Language Learning* 25, 209–235.

Schumann, J. H. 1976. Social distance as a factor in second language acquisition. *Language Learning* 26, 135–143.

Part 4

Organizing and Evaluating Teaching and Learning

Toward A Situational Curriculum

Bernard A. Mohan[1]
University of British Columbia

The curriculum organisation of situational approaches to language teaching is problematic and its relation to a structural syllabus unclear. It is suggested that the task of organising a situational curriculum can be made easier by relating situations to four 'classes': description, sequence, decision-making and generalization; and by relating each class to part of the structural syllabus. Furthermore, the classes have an additional practical value because they connect with semantic notions, speech acts, rhetorical categories and certain teaching techniques.

Many writers distinguish between two major approaches to second language teaching: the structural approach and the situational approach. There is a clear and specific curriculum organisation to the structural approach, but the same cannot be said for the situational approach. Yet the question of the curriculum organisation of the situational approach is important. There is a recent trend to increase the emphasis on the situational component in language courses and those who develop such courses are likely to find themselves facing this question as a pressing concern. The goal of this paper is, therefore, to present an idea for organising the situational curriculum which helps to relate together a number of different elements.

Since the term 'situational' has been used quite loosely in the literature of language teaching, I will specify the aspect that I will be dealing with by contrasting it with some other possibilities. 'Situational' can be understood to include: a goal of language teaching viz. the students' ability to apply their knowledge of L2 to real speech situations (communicative competence); certain classroom learning activities, e.g. role-playing; a way of presenting teaching points (situational presentation); a way of selecting and organising content in a curriculum. This last is my particular concern. In the structural approach the organisation of the curriculum has two facets: selection, or the specification of what is to be learned, e.g. a listing of grammatical structures or a statement of grammatical rules, and gradation, the sequencing of the teaching points drawn from this specification. While I believe that both facets apply to the situational curriculum, I will only be concerned here with the selection facet of the situational curriculum, with the specification of what is to be learned, or the content of the curriculum. This means that it will be left an open question as to how

---

[1] I would like to thank Winnie Au-Yeung, Rick Berwick, Casey Komori, Jill O'Brien, Jim Placzek, Maureen Seesahai, Pat Wakefield, Dan Walsh, Caroline Yi, Nancy Yildiz and Aileen Yip. Their criticism was patient and thoughtful and their suggestions most helpful.

this content might be grouped into units and lessons for a particular level of student.

While the structural approach and the situational approach are different, they do not necessarily exclude each other. Many language teachers use components of both. But current practice, blending both, tends to one or the other of two alternatives: to organise a course around situations and deal with structures as they arise, opportunistically; or to organise a course around structures and to choose situations solely to illustrate structures. The dilemma seems to be that either structures or situations must be dealt with arbitrarily; it is difficult to organise a curriculum so that both are treated systematically. I will therefore be concerned to show how a situational curriculum can be organised so that there can be a predictable relationship with structures.

How is situational course content presently organised? Some current examples are provided by phrase-books and by orientation courses. 'Frases en inglés para defenderse' (Butovsky & McHugh, 1976) is a good example of the phrase-book genre. It is essentially made up of a series of topics (e.g. 'Shopping'). Each topic covers some representative situations (e.g. shopping for food, shopping for clothes), and a situation consists essentially of a list of useful sentences and phrases. (In some other phrase-books, dialogues are given.) The situations chosen are typically oral (requiring both speaking and listening), face-to-face situations oriented to action (i.e. a purchase is made.) 'Orientation Resources' (Butterworth, 1975) is a useful source of information on the organisation of orientation courses. It lists information materials on various topics (e.g. Health, Finances, Consumerism, Tenancy, Your Child's School) for teachers working with immigrant women to use as the basis for lessons. It is suggested that the students' needs and interests are to determine which topics are chosen.

These examples suggest several observations. Firstly, situational course content is not as specific and limited as might be thought, since it does not consist of situations per se, but rather it consists of topics, which are illustrated by situations which show how an interaction in that topic area might develop. Secondly, a selection of these topics can be made on the basis of student need and interest. While this is not as simple as it appears (for a close examination of issues involved, see Mastai (in progress)), it is clear how this could be done in a general way. Finally, since the general approach for the selection of topics is clear, the question of selection in a situational curriculum can be narrowed to the question: given any particular topic, how might the content of that topic be organised in a curriculum?

'Situational activities' are often taken to mean speaking and listening activities only and to exclude reading and writing activities (an unnecessary restriction, in my opinion.) Some ESL curricula in reading and writing have an organisation which is not a structural organisation. Can a similar form of organisation be applied 'across the skills' to the situational material in a topic area? In *Writing as a Thinking Process* (Lawrence, 1972), composition exercises are gathered under 'logical methods or organisation' (e.g. classification, definition) which are largely derived from rhetoric texts. In my view, a number of Lawrence's logical methods of organisation can be grouped into four classes:

Descriptive class: spatial order, comparison and contrast, classification, definition
Sequence class: chronological order, cause and effect
Decision-making class: discussion, proposal, refutation
Generalisation class: generalisation and specifics

The next step, therefore, is to show how these four classes apply to situational material in a topic area. Taking the topic of shopping and an example situation of shopping for clothes represented by a situational dialogue, each class will be dealt with in turn. For each class there will be: a statement of the basis for that class in the elements and relations of typical situations; an example to show how relevant 'logical methods of organisation' can arise in a conversational interaction in the situation; a sketch of a classroom activity which is intended to encourage students to use the class in a creative fashion (it should be easy to see how students could use the class through simple, controlled techniques, e.g. using model dialogues); a listing of structure points, semantic notions and speech acts which might be explored. In this way it should become apparent how each class can be exploited in a situational curriculum.

1. *Example situation*: Shopping for clothes in a department store.

Shopper: Hello.
Clerk: Good afternoon.
Shopper: Please show me some size 15 white shirts.
Clerk: Here you are. (Shows shirts to shopper).
Shopper: (Decides on choice). EITHER: Thank you, I'll take this one (Pays.)
OR. No, I don't like them.
Clerk: Thank you.
Shopper: Goodbye.

This situation relates to what follows because it implicitly contains the basis (see below) for each of the classes: activities (e.g. 'Shows shirts to shopper') and other non-verbal elements which can be described, a sequence of action which can be narrated and analysed, a decision ('Decides on choice') which leads to alternative actions (Either/Or) and which can be explored, and the assumption of background knowledge of how department stores operate. Hence work with the classes can be developed using a situation like this as a starter.

Even a simple situation gives rise to a family of possible dialogues and activities. Since a dialogue is a conversation made up of speech acts, variant dialogues can show speech act variations of greetings and farewells, opening and closing conversations, indirect forms of requests ('Could you show me. . . .'), offers, refusals, acceptance. . . . Since the incident need not flow smoothly, one can explore attention getters ('Excuse me!'), mishearing ('I beg your pardon'), corrections ('These shirts are blue, not white'), apologies ('I'm afraid we don't have any'), to name a few possibilities. For the use of flowcharts to exploit this potential, see Mohan & Katz (1977).

2. *Descriptive class.*

Basis: a typical situation includes participants, actions, objects and a scene or setting. Dramatisation or pictures show these visually. All can be described.

Examples: shopper gives an *identifying description* of the item wanted. Shopper and clerk *compare and contrast* items. Clerk discusses main types of items (*classification*).

Classroom activity: The students reconstruct the description of the situation as seen initially by the shopper and the description as seen initially by the clerk. In this way they create the materials for a role-play.

Structure points: NP + BE + NP / Adj. / Prep. Phrase, There + BE + NP, relative clauses, reduced relatives, adjectives, demonstratives, articles; NP + HAVE + NP, possessive pronouns, genitives; spatial prepositions; comparative sentences, the same as, different from, and . . . too, but.

Semantic notions: existence, attribution and predication, possession, spatial relations, comparison and contrast.

Speech acts: reference, predication.

3. *Sequence class.*

Basis: A typical situation is a series of related events and actions on a time line. If there is no discourse, we have an action chain. With discourse, we have a script. These can be narrated.

Example: The clerk's attention is distracted, and the shopper leaves with the goods, mistakenly thinking that they are paid for. The clerk calls the shopper back and they clear up the mistake, establishing the *chronological order* of events and the *reasons* for the confusion.

Classroom activity: An extension of the above. Shopper is to be accused of shoplifting. Students create the events which led up to this. Then some take the roles of shopper, clerk and floorwalker and the rest act as the supervisor and interview them about what happened. A report on the incident is written, combining the interview information. Simpler activities include narrating the script or action chain and issuing instructions to the shopper or clerk.

Structure points: Prepositions and prepositional phrases of time, cause and purpose (at, before, because of);
adjuncts of time 'when', and time duration (today, afterwards, since), clauses of time, condition, reason or cause, circumstance, purpose, result;

sentence time relaters (first, rest, earlier, later), tenses; quotation, reported speech; imperatives.

Semantic notions: Time relations between events; conditions, cause and reason; relation of events to present time of speaking.

Speech acts: Report, instruction, explanation (in the sense of identifying antecedents or intended consequences)

4. *Decision-making class.*

Basis: A typical situation will include intentional actions. Any such action springs from a decision (although some decisions are more trivial than others) which derives from a decision situation. The decision situation can be stated.

Example: The shopper makes a decision to buy or not to buy. This can be by internal deliberation or can be by *discussion*. A possible choice can be offered by a *proposal* and can be rejected with a *refutation*. Alternatives can be *compared and contrasted* and their *consequences* explored.

Classroom activity: The main decision is whether to buy or not. The students create the decision situation that this springs from. This can be simpler or more complex, depending on the level of the students. A statement of a decision situation is the basis for a problem solving activity.

It is worthwhile to mention some decision situation elements:

a) Problem-statement of the relevant aspects of the situation
b) Alternative courses of action and their consequences
c) A decision on the basis of:
d) the ranking of alternatives in order of preference, on the basis of:
e) judgments on the alternatives, arising from:
f) the values and goals which are involved

Previous semantic and structural points:

comparison and contrast and consequences of alternatives

New structural points:

Problem—what *should* I do?
Alternatives–I *can* do *either* X *or* Y
Ranking—X is *better* than Y
Judgments—I *like* / *dislike* A, A is *too* expensive, A is stylish *enough*.
Values and goals—I *must* / *ought* to / *need* to / *want* to do B.

N.B. The main structural items here are modals. Mention should be made of related constructions such as 'it is possible/necessary to do X.'

Semantic notions: Ability, obligation, permission, necessity, possibility, probability, alternation, intention, evaluation, attitudes, motivation.

Speech acts: appraisals, verdicts, proposals, refutations, advice, recommendations, predictions.

5. *Generalisation class.*

Basis: A typical situation is specific and particular. Hence it is not obvious how generalisations or general statements relate to it. However, recurrent social situations will assume some general knowledge on the part of the participants. Where ESL students might not have this general knowledge, some phrase-books and language courses provide it in the form of orientation information.

Example: Berwick et al. (1977) gives this information for the shopping situation, in *general statements*:

a) Classifying stores (department, discount, specialty) according to quality, price, and range of goods, and type of service.
b) The allowable means of payment (cash, charge, cheques)
c) The procedure if one is dissatisfied with a purchase.

Specific case: shopper wishes to pay by cheque. Clerk states a *general rule* that two pieces of identification are necessary. Showing a personal letter, the shopper asks whether this *specific item* will do.

Classroom activity: Using the general information, the students invent incidents to which the information applies, e.g. choosing the best store for a particular purpose, returning unsatisfactory goods.

Structure points: relating to general statements (specifics have been covered previously.)
Zero article in general noun phrases (shoppers may return goods)
Universal and partitive pronouns and determiners (e.g. everyone, all, some, any, nobody)
Quantifiers (many, few, much)
Time frequency adjuncts (daily, normally, always, sometimes, never)
Summative and appositive sentence connection (generally, for instance)

Semantic notions: quantification, classes or universals, class inclusion.

Speech acts: definitions, summary statements, generalisations, rule statements, explanations (in the sense of giving a general statement which subsumes a particular case.)

The examples above show that situational materials in a topic area can be organised under the classes of description, sequence relations, decision-making and generalisation. Only one topic was used, but this was for purposes of illustration only, in order to be specific and to make clear the links between the elements in the curriculum, for my assumption is that these classes will apply to situational materials in any topic area. Earlier it was said that it was difficult to organise a curriculum so that both situations and structures were treated systematically. My solution has been, not to attempt to relate situations and structures directly, but to relate them via the classes: to claim that the classes can be related to the situations in any topic area and to claim that each class can be related to a particular group of structures. This opens up the possibility of working through situations and structures simultaneously, and the classes emerge as organising centres for a situational curriculum.

As was said above, I have left open the question of how the content of a situational curriculum might be developed into a graded sequence of units and lessons, but a few comments are in order. In an actual situational curriculum a range of topics would be chosen and the classes could be explored in a number of different situations. It is possible that description would occur early, sequence relations later and the more complex aspects of decision-making later still, and that generalisation would enter at a number of points. In addition, there would probably be a cyclic development of the classes, so that the gradation would spiral through situations, treating the classes in increasing depth of difficulty. Clearly, the issue of gradation in a situational curriculum is more problematic than sequencing a number of structural teaching points, and will require further work.

Support for the claim that the classes can be related to the situations in any topic area is provided by the example section and the basis section for each class. The example section merely illustrates how components of a class can be used naturally in situational interaction. The basis section argues that each class relates to a typical aspect of situations. All of these aspects are based on the idea of social action—its elements, its antecedents and consequences, the decision situation it emerges from and the norms and procedures it corresponds with. The assumption is that social action is fundamental to situations and that its characteristics form a principled basis for the classes. (This raises the possibility that there may well be more than four classes. But four appear to give a sufficiently complex framework to begin with). A standard objection to the concept of 'situation' is that it is either simplistic or incoherent and in either case is an unworkable foundation for a curriculum. I have suggested that the concept has a structured complexity that can be exploited in curriculum design.

What is the practical value of the classes as organising centres in a curriculum? One point is that they can be related not only to particular groups of structures but also to semantic notions and speech acts. The listings of these that I have given are tentative, speculative but (I hope) plausible. An essential next step will be to check them by developing and testing classroom materials. Another point is that the classes show a natural relation to different ways of stimulating talk in the classroom. Each of the classroom activities suggests this. Description relates to ways of using visual material, most obviously in describing pictures. Sequence is simply developed in strip stories and in more complex forms of narration: the classroom activity given for sequence is based on a 'project method' technique in which the students create 'soap-opera' material (see Central Committee on English (1970) for a full description of it as an L1 technique and see Morrow & Shaw (1974) for its extension to L2 learners.) Decision-making is most clearly involved in problem-solving activities currently familiar in ESL teaching (cf. Kettering, 1975) but its full scope is best seen in values clarification techniques (cf. Howe & Howe, 1975). While generalisation connects with methods which match principles (or rules, or procedures) with examples, and such methods are used throughout teaching (cf. law, mathematics), it is difficult to find any use of them to promote language activity specifically. A final point is that the classes parallel the rhetorical categories of description, narration, argument and exposition. The classes apply to situational material while the rhetorical categories apply to extended written monologue. The usefulness of this parallel is that it forms a potential bridge between a situational curriculum and more advanced and academic courses in reading and writing which are developed around the rhetorical categories (e.g. Lawrence, 1972).

It will be clear that I have not provided a situational curriculum. Instead I have put forward a framework of four classes which should help to simplify the complexity of the task by acting as organising centres to connect a variety of different elements in a curriculum. The real test of these ideas is whether they prove useful to those developing materials for the classroom.

REFERENCES

Berwick, R., M. Blackwell, L. Der, and S. Sair. 1977. *Practical English revision*. Vancouver, Department of English Education. University of British Columbia. (Unpublished mimeograph).

Butovsky, L., and J. McHugh. 1976. *Frases en inglés para defenderse*. Toronto, Ontario Ministry of Culture and Recreation.

Butterworth, B. (ed.) 1975. Orientation resources. *T.E.S.L. Talk 6*, 2, 1–222.

Central Committee on English, 1970. *Projects in practice*. Edinburgh, Her Majesty's Stationery Office.

Howe, L., and M. Howe, 1975. *Personalising education: values clarification and beyond*. New York, Hart Publishing Co., Inc.

Kettering, J. 1975. *Developing communicative competence: interaction activities in English as a second language*. Pittsburgh, The University Center for International Studies.

Lawrence, M. 1972. *Writing as a thinking process*. Ann Arbor, The University of Michigan Press.

Mastai, J. (in progress) Immigrant learning tasks. Ph.D. thesis. Vancouver, Department of Adult Education, University of British Columbia.

Mohan, B., and N. Katz. 1977. Communicative competence and flowcharts. *T.E.S.L. Talk 8*, 1, 25–31.

Morrow, K., and M. Shaw. 1974. L_1 techniques in E.S.L. teaching. *English Language Teaching 28*, 3, 197–202.

Parameters of Language Syllabus Construction

William E. Rutherford

American Language Institute
University of Southern California

There has been a long standing opposition in second language pedagogy between those who take (often implicitly) as the fundamental point of departure for syllabus construction *language organization* (e.g. Rosenbaum, 1964; Ritchie, 1967; Robinett, 1973; Berman, 1975) and those who take *communicative function* (e.g. Newmark & Reibel, 1968; Currie, 1975; Levine, 1976). In a sense, both positions are right, in that some students tend to learn better through focus upon formal aspects of the language, others through focus upon the use of the language. Since typical classes always include varying mixtures of the two, it makes sense to let the course syllabus be constituted such that it meets the needs of both kinds of learners. Ministering to both, however, requires that the syllabus compiler be aware of the effects attention to language form and attention to language pragmatics can have upon each other. For example, presenting the facts of verb and noun complementation (i.e. strict subcategorical restrictions) will force the introduction of large amounts of vocabulary, much of which can not, for all practical purposes, be treated in context. Again, use of the language for whatever function will point to grammatical choices among semantically equivalent structures which cannot be made unless the structures are already a part of the user's knowledge. Other interactions touch on such concepts as functional need, rule accessibility, rule breaking, error, and grammatical choice. More understanding of such interpenetrations can contribute significantly to the improvement of language syllabus construction.

Exactly one decade ago, when transformational linguistics was at its most monolithic and its adherents at their most vociferous, there appeared in the journal *Language Learning* an article (Ritchie, 1967) discussing the implications of generative grammar for the compilation of ESL syllabuses. The discussion focused essentially upon the answer to the question, "Given a linguistic description of the language to be learned, what tasks are to be assigned to the learner which will provide him with the best opportunity to acquire the information embodied in that description?" (p. 111). The "information" here is of course structural information, for a few pages later we read that "a course will be based first and foremost on the syntactic structure of the foreign language" (p. 120). Semantic considerations were limited to choice of lexical items to be plugged into the syntactic framework at any given point and phonological considerations consisted of pronunciation practice. This arrangement was what one might term an extreme version of syntactically oriented syllabus construction.

Such orientation has also been suggested before this time (e.g. Rosenbaum, 1964) and since then has been assumed in papers by, among others, Robinett (1973). Berman (1975), and Starosta (1976), as well as myself (Rutherford, 1968). Furthermore, the period of the last decade or so has been one where a number of pedagogical grammarians have seen fit to publish an article or two—TESOL Quarterly contains quite a few of these—which attempt to adapt for classroom purposes the latest collection of insights from the major linguistic journals. It was usually assumed that if linguistics had discovered something new about English, then it was someone's duty to publish a simplified version of it for language teaching. More often than not, however, it has been left up to the teacher to decide what use, if any, to make of these "insights."

The extreme version of syntactically oriented syllabus construction never came close to dominating language pedagogy, however, even at the high point of transformational-generative ascendancy. Other, long-standing language teaching traditions were never affected by it and college and university ESL programs opted as much for materials with a situation base as for those with a structure base. Grammar itself was seldom lost sight of, however, and we will return presently to the ways for pedagogy in which language form and language function can interact.

The European tradition of syllabus construction has usually taken communicative function as its point of orientation (Currie, 1975; Levine, 1976), an approach supported in this country by, among others, Newmark and Reibel (1968). In the extreme version of the functional approach, where models of language in use are all that is required, attention to language form is not a feature of syllabus construction. The role of linguistics in these two extremes is thus paramount in one and negligible in the other.[1] Yet, it is obvious that even the extremes cannot totally exclude each other. For example, the structure based program has to involve at least classroom communication of the tasks to be performed and the situation based program will have to handle in some graceful way the inevitable classroom questions about language form. But leaving the extremes aside, we are still left with a great deal of room for maneuvering. Attention to form and attention to function can affect the classroom, and be affected by it, in a variety of ways. The variables that can play a part include the tendency for individual students for whatever reasons to feel more comfortable with the emphasis one way or the other and hence to learn faster that way; the tendency for at least comprehensive programs to be "compartmentalized" in such a way that a form/function dichotomy is implicit in the very rubric of compartmentalization (at our own American Language Institute, for example, classes labeled "spoken English," classes labeled "structure"); the frequent expectations of students—usually those with an academic background—that classroom activity will concern itself with matters of form, and bewilderment, bordering occasionally on hostility, upon realizing that that's not always the way we do it.[2] Ministering simultaneously to the needs of students with different learning biases, however, requires that the syllabus compiler be aware of the effects attention to language form and attention to language function can have upon each other.

---

[1] One of the healthiest characteristics of the functional approach, however, has been the tendency to look for linguistics for help only when a problem arose which seemed amenable to linguistic solution, rather than letting a linguistic analysis serve as the rationale for construction of the entire syllabus.

One of the dangers of intense focus upon structure is that language comes out sounding something less than natural. Unnaturalness in general can be of various kinds (e.g. incompatible registers: Hi there, your Excellency; illocutionary mismatch: Could you tell me what time it is? Yes, I could; artificial redundandancy: Is your name Rick? Yes, my name is Rick; etc.). But the greatest danger of concocting sentences that uniformly, even ruthlessly, display some selected syntactic feature is the kind of unnaturalness that arises because the information-bearing function served by that feature has been lost sight of. Offenses of this kind often occur when a teacher is hard-pressed to devise an exercise focusing on a specific point of structure to be covered in class that very day. Let's say that for one reason or another it happens to be relative clause day. Nothing could be simpler than creating two sentences having a shared noun phrase and letting one be relativized within the other. So on the board, or in a handout, appear for student enlightenment museum pieces like "The boy who bought a chicken for his grandmother doesn't like peanut butter" or "The bicycle rider struck the salamander whose tail the cat bit off." What makes such sentences still bizarre even with an invented context is that the relative clauses don't seem to be doing anything, performing any function. The linguist Susumu Kuno, in a recent study (Kuno, 1976), has shown that a relative clause must be a statement about its head noun and that violation of this principle automatically explains some of the deviance which previously had to be accounted for through the so-called coordinate structure contraint (Ross, 1967). Although the example sentences I cited come close to being deviant in this way, total unnacceptability is the only label for an example like this one, where a comparative construction has been relativized:

**The bicycle rider struck the little boy who Mary is taller than.*

(This example and others like it were recently used as part of a test to determine non-native speaker comprehension of relative clauses.) Here the relative clause is in no sense a statement about the little boy; yet in a purely formal sense the sentence would have to be judged grammatical. There is a matrix S (The bicycle rider struck the little boy), a constituent S (Mary is taller than the little boy), and a shared noun phrase (the little boy), and for purposes of sentence concoction these conditions are all that is required. Kuno explains, however, that when clauses like this occur we have what he calls a conflict of empathy focus. We are led to empathize with "the little boy," the shared NP in the main clause, but are forced to empathize simultaneously with "Mary," the subject but not the shared NP of the relative clause. Hence the oddness. My point is that with preoccupation with language form for whatever purpose, it is easier than one may think to fabricate language examples which do not just ignore but actually violate principles of languages function. Possible ways to avoid such a pitfall would be to elicit other native-speaker intuitions about the items intended for classroom use, or to supply the items with some kind of context, or better still to lift samples from actual language, all of which of course becomes increasingly time consuming.

---

[2] Even Newmark (1968), in arguing against structurally based syllabuses, cites the students' "craving" for formal generalizations.

One of the commonest language terms in general use is the word "grammar." Now "grammar" in no sense means the same thing to all people. To a school child, grammar is likely to be viewed as a proscription device most frequently brought to consciousness by the familiar adult admonition "Watch your grammar." To the American public school system, grammar usually represents a collection of informally stated, empirically unverifiable "rules" the learning of which is designed to bring about student speech and writing patterns more in line with establishment norms. To the faculty at many contemporary linguistics departments. "grammar" signifies the integrated systems of syntax, lexis, and morphology, and to the transformationalist grammar will be synonymous with "language" itself.

The term "grammar" is no less familiar to the world of language pedagogy, as we all know. And there are various ways of subdividing the substance of pedagogical grammar: frequency of invocation, amount of required mental baggage, communicative function, amenity to intra-lingual explanation, etc. What I would like to call attention to, however, is a distinction between, for lack of better terminology, lexical and non-lexical grammar. By this I mean grammar the learning of which is tied to the learning of large amounts of vocabulary, and grammar whose principles can be grasped independently of a specified lexicon. Lexical grammar focuses upon, among other things, those lexical items (i.e. nouns, verbs, adjectives) that in Chomsky (1965) carry a feature indicating application of a rule of strict subcategorization. Non-lexical grammar, which embraces many of the major syntactic processes of English, is not tied to a specific lexicon. It is not necessary, for example, that the shared noun phrase of matrix and constituent be of a certain class or type in order for a relative clause to be grammatical. Relativization, therefore, is an example of non-lexical grammar. Non-lexical grammar also includes such areas as negation, interrogation, comparison, concatenation, the tense system, etc. Lexical grammar takes in complementation with its strict subcategorial restrictions on verbs, violations of which give rise to errors like **I avoid to smoke*, **I wonder that how is it possible*, **I hope coming again*, etc. Other syntax that falls within lexical grammar are the phrasal verb system with its long lists of verb + Particle and verb + preposition, derivational morphology with its long lists of noun-verb-adjective correspondences (expect ~ expectation ~ expectant, prefer ~ preference ~ preferential), and topicalization, a subdivision of which—active vs. passive—is built upon one property of a large number of verbs, viz, transitivity.

To attempt to teach the principles of these various instances of lexical grammar would necessarily inundate the student with massive amounts of associated vocabulary, which for all practical purposes cannot be treated in context. I'am reminded of those structure-based texts that "wrap up" the phrasal verbs in a single unit or that lay out a large group of sentence-combining exercises utilizing the rules, say, of English complementation.

Complementation and its associated vocabulary will need to be spaced out, doled out a little at a time, so that each verb, noun or adjective + complement can be looked at together. Another reason for proceding in this way is that the selection of complementizer is largely a principled one and the matter of the learner's choosing correctly need not in any case depend entirely upon his remembering some otherwise arbitrary strict subcategorization restrictions on long lists of verbs, nouns, and adjectives, since, as shown by the Kiparsky's (1968), these restrictions can to some extent be predicted from the meaning of the lexical items themselves. The question of whether many more syntactic phenomena are actually controlled by non-syntactic factors is being given more and more attention these days. Kuno, at the end of his relative clause article (1976) cited earlier, writes that "it is time to reexamine every major 'syntactic' process and every major 'syntactic' constraint from a functional point of view, to find semantic explanations for its existence in case the syntactic characterization holds, and to find a deeper and more accurate semantic generalization in case the syntactic facts are simply superficial and 'almost correct' syntactic manifestations of nonsyntactic factors." (p. 438)

Structural generalizations for pedagogical purposes don't always have to be couched in structural terms, much less linguistic jargon. There is a split, for instance, of ambiguous English indefinites (e.g. a handkerchief) into specific (I'm looking for *it*) and non-specific (I'm looking for *one*). One very effective way of calling attention to this—at least at the intermediate level—without requiring any mental gymnastics is to tell the story of the famous demographer who gave a lecture on the world population explosion. In order to emphasize the seriousness of the situation he pointed out to his audience that somewhere in the world a woman gives birth to a child every thirty seconds. With that a man in the audience rose with clenched fist and shouted "We must stop that woman!"

Some of the cartoon strips in the daily newspaper frequently derive their punch through exploitation of a fact of syntax. Noun phrase reference deletion was the focus one day for the "Crock" strip, which always has a Foreign Legion type setting. In the first two frames a corporal hands pencil and paper to a legionnaire prisoner in solitary and says "The commandant wants a written apology for throwing his cat in the mashed potatoes." In the third frame the prisoner starts writing "Dear cat . . ."

The cartoon and story devices just mentioned highlight the application of rules which are relatively formal and somewhat complex, rules which would constitute somewhat cumbersome "mental baggage" for the student to carry around in his head. As such, these devices would seem to mark small areas where structure and situation nearly overlap, where attention to one need not occur at the expense of the other.

Occasionally the cumbersome rule or rules—and these are usually real linguistic rules—can be restated, at least in part, as a "pedagogical" rule, what I call a "rule of thumb." Rules of thumb seem to be formulated most often as statements of co-occurrence, which are perhaps the least difficult kind to remember. In this category I would include not only the various kinds of rules of agreement and concord but also those more aptly expressable as relationships and compatibilities: the co-occurrence requirements of certain time adverbials with certain verb tenses, for instance. One example of an extremely simple rule of thumb extractable from an immensely complex bit of grammar has to do with the recognition of such determiner errors as "Who has telephone?" and "I have to read book." The rule is that a count noun in the singular cannot stand by itself. It has to occur with a determiner: "a, the, my, your, this, that, one," . . . any of these, . . . but *something*. Application of the rule doesn't guarantee that an appropriate determiner will be chosen, but it will at least ensure that the noun does not go unaccompanied, which is one of the commonest errors made by learners of English whose native languages are non-Indo-European.

The applications of many rules serve some kind of communicative purpose, as is well known. It also happens that the calculated *mis*-application of a rule can also serve a purpose, and we see this most often in the world of graphic advertising, where deliberate rule-breaking is designed to do nothing more than catch the eye and attract attention. ("Suddenly the fastest way to Europe is also the comfortablest."—Pam Am). One of the most fascinating ways of focusing classroom attention on language form while developing a situation which is far from artificial or contrived is to have the class take a close look at the ungrammatical advertising that has a method to its madness. Let me turn back again briefly to NP-determiner co-occurrence. Some time ago "Leo's" luncheon meats displayed a full page magazine ad showing a sandwich a foot high and filled with Leo's products. The caption at the bottom read "Leo's chicken and turkey makes a great sandwiches." In the students' animated discussions about what idea was supposed to be conveyed by purposely wrenching the internal structure of the noun phrase, unconscious attention was drawn to the *correct* co-occurrence rules and the students reinforced their understanding of several basic determiner rules in English without it being driven home in some more heavy-handed fashion.

Choice of form is vacuous if it serves no communicative purpose and function is a meaningless concept without the formal means to implement it. From the standpoint of syllabus compilation then, the area where form and function interact most closely is what Widdowson has labeled "usage," if I understand his term correctly. Wallace Chafe (1976:28) would call it "packaging," the arrangement of language content in such a way that the hearer or reader can process it with a minimum of effort or difficulty. However, choice of packaging for a student presupposes that the forms of the choices are part of his knowledge of English.

Now the focus of classroom attention upon alternate ways of saying or writing things is nothing new. In textbooks new and old, exercises abound in which the student is required to say or write a sentence in some semantically equivalent but syntactically different version. The term often used in the instructions for such exercises is "transform." Courses are frequently laid out in such a way that the student is "exposed" first to the so-called "basic forms, then to the various ways in which such forms can be, as it were, "transformed" or combined. After an ample amount of sentence-level work of this kind the activity that usually follows is that in which the student freely expresses himself, whether in speech or in writing. But this free expression, even if grammatical violation is kept to a minimum, invariably produces a kind of jarring effect, where individual sentences appear awkward not singly but only in the context within which they've been placed. The student has been taught different ways of saying or writing things, so it is not necessarily the case that he doesn't know it's possible for an English sentence to assume a certain shape in a particular context; it's that he usually doesn't know that there are *reasons* for preferring one shape over another, in that context or any other.

One of the best informal statements of the criteria by which grammatical constructions are deemed appropriate or not is to be found in the last chapter of that monumental volume by Quirk, Greenbaum, Leech, and Svartvik, *A Grammar of Contemporary English* (1972). Very briefly, those grammarians point out that there is a general tendency in English to let "old" or given information come early in a sentence and "new" or salient information come late, the so-called principle of "end focus." Parallel to this is another tendency to put "complex" constructions last in a sentence, the so-called principle or "end weight." Furthermore, since what is complex—especially, for example, surface vestiges of underlying sentences—is extremely likely also to be new information, the two principles work with rather than against each other. Suppose, for instance, one had just written the sentence "Kinesiology serves a useful purpose" and the next idea was to be a statement of that purpose. From the standpoint solely of content, that purpose might be expressed either as:

> In order to learn how to analyse the movements of the human body we study it.

or:

> We study it in order to learn how to analyse the movements of the human body.

The criteria of usage just mentioned would of course require that one express the idea by choosing the second alternative rather than the first.[3]

---

[3] Notice here also how a principle of discourse automatically explains a factor of sentence arrangement that would otherwise have to be reduced to grammatical terms: that generally speaking anaphoric elements make for "weak" endings of sentences. Since by definition anaphora cannot represent new information, it will usually not occur at the end of the sentence.

For language teaching materials then which feature the bringing of matters of language structure to the grammatical consciousness of the learner, it is not enough simply to present to the student the various syntactic forms in which any particular meaning can be realized (active vs. passive, for example); it is necessary also for the student to know that in practically any linguistic environment the choice among these related grammatical forms is not arbitrary, and he should know enough of the principles of sentence arrangement (i.e. usage) to be able to make these structural means appropriate to his rhetorical ends (i.e. use). The materials writer needs to be aware, however, of how the end and the means can affect each other, because the extent of that awareness will be reflected in the quality of his product.

REFERENCES

Berman, R. 1975. Analytic syntax: a technique for advanced reading. *TESOL Quarterly* 9, 3, 243–251.

Chafe, W. L. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and points of view. In Charles Li (ed.), *Subject and topic*, New York, Academic Press.

Chomsky, N. 1965. *Aspects of the theory of Syntax*. Cambridge, Mass., MIT Press.

Currie, W. B. 1975. European syllabuses in English as a foreign language. *Language Learning* 25, 2, 339–354.

Kiparsky, P., and C. Kiparsky. 1968. Fact. In Bierwisch and Heidolph (eds.), *Recent Advances in Linguistics*, The Hague, Mouton.

Kuno, S. 1976. Subject, Theme, and the Speaker's Empathy—A Reexamination of Relativization Phenomena. In Charles Li (ed.) *Subject and Topic*, New York, Academic Press.

Levine, J. 1976. Some socio-linguistic parameters for analysis of language-learning materials. *IRAL 14*, 2.

Newmark, L., and D. A. Reibel. 1968. Necessity and sufficiency in language learning. In Lester (ed.), *Readings in applied transformational grammar*, 2nd ed., New York, Holt, Rinehart & Winston, 1973. Pp. 220–244.

Quirk, R., S. Greenbaum, G. Leech, and J. Svartvik. 1972. *A grammar of contemporary English*. Longman, Inc.

Ritchie, W. C. 1967. Some implications of generative grammar for the construction of courses in English as a foreign language. *Language Learning* 17, 3–4, 111–131.

Robinett, B. W. 1973. The application of linguistics to TESOL: once more. *TESOL Quarterly* 7, 4, 425–435.

Rosenbaum, P. S. 1966. On the role of linguistics in the teaching of English. In Janet Emig *et al.* (eds.) *Language and learning*. New York, Harcourt, Brace & World.

Ross, J. E. 1967. *Constraints on Variables in Syntax*. MIT Doctoral Dissertation.

Rutherford, W. E. 1968. Deep and surface structure and the language drill. *TESOL Quarterly* 2, 2, 71–79.

Starosta, S. 1976. A place for case. *Language Learning* 26, 1, 1–36.

Causes of Failure and Conditions For Success in the Learning and Teaching of Foreign Languages

Peter Strevens

Cambridge University

Learners are more than walking language acquisition devices: the extent of their success in learning a foreign language depends on achieving the maximum fit between many interlocking variables, some more important than others. The roots of these variables lie in some cases in society, in other cases in the teaching profession, in others within the individual learner. 'Success' or 'failure' must be seen in relation to particular learning/teaching aims, but some factors can be identified as restrictions on success, or as conditions that maximise it, *in general*.

Among the strongest reasons for failure are: unwillingness to learn; certain physical and administrative impediments; insufficient time for learning and teaching; unrealistic aims; a false relationship between the nature of teaching materials and the professional standards of the teachers who will use them; gross incompetence in the management of learning. Less seriously, learner expectations may be low; teachers may be inadequately prepared in various ways.

Success is most likely to be achieved with willing learners; when physical and organisational impediments are removed; when teachers are appropriately trained. An analogy will be drawn with Intensive Care Units in medicine. The foregoing analysis will be applied to some different types of foreign language learning/teaching situations.

This concept of success and failure derives both from studies of 'the whole learner' and from a model of the learning/teaching process. Achievement is highest when skilled teachers are encouraged by society and their profession to cherish willing learners.

A characteristic of the language teaching profession is its enormous diversity and variability. Teachers of languages throughout the world are numbered in millions: learners of languages are counted in scores of millions. Hundreds of languages are involved, either as the target of learning or as the mother tongue of the learner. There are great numbers of different aims and objectives, different rates of intensity, different methods and materials and styles of learning, different levels of proficiency aimed at. And there is an equally wide range of results, from total failure to learn, to rapid and easy achievement.

Attempts in the past to overcome the variability of learners' achievement—that is, to counter the lack of success of individuals or groups—have tended to concentrate on finding blanket solutions. We have justified our changes of direction by saying such things as: 'The method must be wrong'; or 'We need contrastive analysis': or 'We ought to use language labs'; or 'We ought to stop using language labs'; or 'We must rely on linguistic theory'; or 'We must change our allegiance from one linguistic theory to another'; or 'We must change our classroom organisation'; and so forth. Yet none of these changes brings about the overall improvement in learning that we assume it should produce, even though *some* improvement is often observed in *some* cases.

The purpose of this paper is to suggest that single, blanket solutions cannot be achieved. A close analysis of reasons for failure to learn, and of the conditions in which success is most often obtained, reminds us that while all human beings can be regarded as walking language acquisition devices and to that extent can be assumed to possess a universal potentiality for learning languages, their success in actually doing so in a framework of organised learning and teaching depends on very much more than this: it depends upon achieving the maximum harmony between a large number of variables, whose precise importance differs from one set of learning/teaching conditions to another.

In discussing the problem I shall begin by briefly considering the nature of 'failure' and 'success', after which I shall suggest that a number of strong recurrent reasons for failure in language learning can be identified, as well as a number of strong recurrent conditions for success. I shall conclude by suggesting that these sets of factors have more than conversational or anecdotal significance because they derive from a theoretical model of the language learning/teaching process.

The Nature of 'Failure' and 'Success' in Language Learning

Used in a general way, the terms 'failure' and 'success' are imprecise and emotive; they gain force by being related to specific aims. The more precisely we can state the targets of achievement for specific courses of learning and instruction, the more accuracy we can give to these terms 'failure' and 'success'. For the sake of argument I shall assume in what follows that teaching is always given a precisely-specified set of aims and targets, and that what we are referring to is the extent to which these are attained in the time devoted to them. Furthermore, since 'failure' can be regarded as 'negative achievement' and 'success' as 'positive achievement', I shall use the term *achievement* henceforth as the general label.

There are, in fact, three different kinds of failure and success: the achievement of the *learner*, the achievement of the *teacher*, and the achievement of the *system* within the learner and the teacher come together.

The achievement of the learner is the extent to which (a) he reaches his potential optimum rate of learning, and (b) he achieves control of the particular linguistic, functional and communicative devices that make up the aims of his learning programme. Normally, language teachers have concentrated on the second of these sub-divisions, i.e. on the learner's terminal command of the sub-set of the language he was seeking to learn. But in a deeper sense the learning/teaching process seeks (or should seek) not only to implant the terminal aims within the learner but also to activate his capacity to learn, and to do so at his *optimum learning rate*—that is, at the fastest rate that he can sustain for the required period of learning.

The achievement of teachers is often measured by their record of examination passes among their pupils. This may be a convenient public index of pedagogical success, but in relation to the fundamental processes at work, examination passes are rather trivial. More importantly, the achievement of the teacher is the extent to which he or she achieves the *optimum management of learning*. To consider teaching as 'the management of learning' enables us to incorporate under a single concept both the learner's and the teacher's contribution to the learning process. This definition also accepts any and every approach, method and technique: it is neutral and non-specific as to methodology.

The achievement of the system is the extent to which it permits and even *encourages* the learner and teacher to reach their best rate of achievement in their respective roles. The 'system' includes (a) the sociolinguistic consensus—such factors as the popularity of a particular language, and the social status of learning and teaching languages—(b) the efficiency of the organisation and administration of education, and (c) the particular school or other framework (including arrangements made for evaluation and assessment) within which the teacher and the learner come together.

It is my strong belief that each of these three kinds of achievement has an indispensable role to play in the total learning/teaching process. Let us consider an example of each: an unwilling *learner* can cause failure in his own achievement, in spite of a good teacher and a harmonious system, though equally an energetic and sophisticated learner can sometimes achieve success in spite of a poor teacher or a low-grade system; an incompetent *teacher* can prevent a learner from achieving success, but equally a good teacher may overcome learner-resistance or organisational faults in the system; a poor *system* can frustrate the efforts of learner and teacher alike (for instance, by allocating insufficient lesson time) but equally a well-designed system can often compensate for the less well-trained teachers or indifferent learners.

To summarize: Success and failure can be re-stated as positive and negative achievement; they are best considered relative to specific aims; they relate in part to the learner, in part to the teacher, and in part to the system within which the learning and teaching occur.

Given this frame of reference for discussing achievement, we can ask a crucial question: whether any particular circumstances are frequently associated with failure or success. I believe that two sets of such factors can be identified, and these will now be described.

Strong Recurrent Reasons for Failure

1. *Unwillingness to learn*. The attitude of the learner towards the task of learning spans a range from total unwillingness through passive neutrality to positive enthusiasm. There are many different reasons—personal, social, psy-

chological, etc.—for a learner's state of willingness, and a skilful teacher can often increase its extent. But the *un*willing learner usually fails to learn.

2. *Learner's expectations are too low*. Many learners, especially adolescents and adults, come to language classes convinced that they will not succeed. The reasons for a negative self-view are many, and often they can be modified or changed into high expectations by a skilful teacher; but low expectations are a serious impediment. I have recent experience of two examples: (i) a university student whose course required her to learn Spanish informed us that she was unable to learn languages. In spite of patient and caring teachers, her view of herself remained inviolate for several months: she learned virtually no Spanish. Then suddenly after about five months she started to make good progress. But she found it so difficult to accept this falsification of her predictions about herself that she preferred to withdraw from the course rather than go on learning Spanish.

A major French oil tanker company is sometimes required to send a captain to take over a ship with an English-speaking crew, at short notice. Courses of up to 100 hours are arranged, but these non-academic learners do not believe that they can 'learn English' in that time, so they begin the course with expectations of total lack of success. Except in extreme cases, the learners' expectations of their own success are based very closely on the opinions expressed by their teachers. It is therefore possible by careless talk to blight the achievement of a learner; equally it is possible by encouragement to bring about considerable improvement.

3. *Unrealistic aims*. These are generally a fault of the system rather than of the learner or the teacher. A glaring example of this factor is the commercial language school that advertises 'You can speak English in 24 hours'. But many school language programmes make presumptions about the learners' terminal achievements which, as teachers themselves often acknowledge, cannot in fact be reached by the average pupil. The learners generally know this—students are very sensitive about such matters—and thus unrealistic statements of aims on the part of the teacher or the system trigger low expectations on the part of the learner, thereby turning them into self-fulfilling prophecies.

4. *Off-set teaching*. This is a special case of unrealistic aims. It occurs, not infrequently, when learners are expected to *learn* the language but in fact the teachers are required to *teach* something different. Three examples are; (i) when *literature* is taught in the belief that this will thereby teach the language but before the learners have the necessary command of language to follow and understand; (ii) when *linguistics* is taught as a substitute for teaching a command of the language; (iii) when a largely oral English course is required to be taught by teachers who have no effective command of the spoken language.

5. *Physical and organisational impediments*. Obvious examples are: *fatigue*, *heat* or *cold*, *noise* or *distraction*. Weary learners learn little; weary teachers teach little. Great heat or cold or other physical discomforts reduce the rate of learning; excessive noise, whether from other classes, or city traffic, or monsoon

rain drumming on corrugated iron roofs, makes language learning nearly impossible; the classroom which overlooks a swimming pool, or is the constant target for the Principal's visitors, detracts the learners. Other examples include *overcrowding, lack of premises, books or equipment, examination neurosis, absenteeism* on the part of the learner or the teacher.

It is essential to note that all these impediments are ultimately avoidable. Unfortunately teachers are inclined to accept such impediments as acts of God, as features of the universe which are to be borne with fortitude but are not susceptible of improvement by human effort. But in fact they are nearly always the man-made consequences of poor organisation or educational policy. In some ways language teachers are more inclined than others to put up with these impediments. Our colleagues who teach science, or woodwork or other crafts do not tolerate conditions in which they are prevented from teaching their subject: and neither should we. Very often these difficulties are tolerated for far too long by teachers who fail to realise that it is an essential part of a teacher's duty to campaign, individually and collectively, for the removal of such impediments. The campaign may take a very long time, but it should never be abandoned. In these matters the teachers know best: it is essential to their professional integrity that they should never acquiesce, except temporarily, in the continuance of physical and organisational impediments to learning and teaching.

6. *Insufficient time for learning and teaching*. This deficiency relates to the *total quantity* of time for organised instruction, and its rate of intensity, as well as the overall duration of courses. Some courses, particularly 'service' language courses—"English for the engineering students" for example—are allocated by course administrators ludicrously few total hours of instruction. Or courses are persistently taught so thinly, say 3 or 4 hours per week, that learning is almost guaranteed to be equally thin and inadequate. At the other extreme, some courses with a rather small syllabus content are taught over periods of 6, 8, even 10 years: teachers are usually well aware that after the first two years of school instruction in a language there is a strong tendency for standards to decline, not to improve, with all its consequent disillusion and frustration for teachers and learners alike.

7. *Gross incompetence in teaching*. Although standards of entry to teaching are improving, and standards of teacher training are improving even more rapidly, there still exist numbers of language teachers whose management of their pupils' learning is incompetent. The teacher whose personality 'turns off' the learners; the teacher with an inadequate command of the foreign language; the teacher without the necessary instructional techniques—these are some of the most obvious examples. These are defects of the teacher, but also of the educational system that tolerates them, and the teaching profession should continue to campaign for their eradication.

8. *When the teachers/materials equation is not solved.* The higher the general standard of the teachers, the less important are the course-books and

other teaching materials. But when the standard of teachers is low, as it inevitably is for a time in conditions of rapid expansion of education, then the standard of teaching materials assumes great importance. This is what is meant by 'the teachers/materials equation'. Insufficiently-trained teachers working with poor materials face a considerable handicap. When this occurs it is a defect of the system, for whose remedy a three-way collaboration is necessary, between publishers, administrators and teachers.

9. *Teachers inadequately prepared*. As the standards of initial teacher training continue to rise, world-wide, the crux of this problem falls increasingly upon the older teachers who may not have profited from present-day standards. The problem is especially acute when a major change takes place, in approach or methodology, which the teachers who are products of earlier stages of development may not be aware of—or may not accept. An example of this has recently occurred in Britain, where the introduction of French in the *primary* school has faced severe difficulties in some areas because of the unwillingness of some *secondary* school teachers to adjust to the change in their intake of pupils. Here it is the system which may need to persuade and re-educate the teachers in order to permit the continued achievement of the learners.

These, then, seem to me to be the most frequent strong recurrent reasons for failure. All of them are avoidable, though some require major changes of administrative policy. Now let us turn to strong recurrent reasons for success: i.e. conditions which seem regularly to be associated with higher levels of achievement.

Strong Recurrent Reasons for Success

1. *Willing learners*. Enthusiasm for learning is a great asset, and one which teachers can nurture in their learners. Most individuals vary from time to time in their degree of willingness, often as a result of emotional ups and downs in their home life. It is a central component of good teaching that the teacher should be aware of these changes, sympathetic to their causes, and continually seeking to cajole and encourage the learner upwards to his optimum willingness to learn.

2. *Learners see the relevance of their learning*. It is only for a fairly short period in childhood that human beings willingly learn without caring why they should do so. It may be that one of the inherent difficulties in school language learning is that as they approach adolescence children cease to learn insouciantly, without caring or knowing why, and start to expect clear and valid reasons for continuing to give their willingness to learn a language. Rates of achievement are greatly improved when learners know and accept the reasons for learning.

3. *Learners' expectations are high*. A belief that one *will* learn is a powerful incentive to actually do so.

4. *The target language has good standing in the community*. It is a sociolinguistic truism that learning and teaching an unpopular language is a difficult task. By contrast, when the foreign language has a high status or is taken for

granted, achievement tends to be high. (This is a variable over which the teacher has little control, as learners' attitudes tend to be established before they start to learn, along the lines of current popular opinion and prejudice.)

5. *Physical and organisational requirement are met.* Two different types of provisions have to be considered. The first relates to the absence of the impediments noted earlier: fatigue, extremes of temperature, etc. The second concerns the actual provision and the efficient working of all necessary buildings, plant, equipment and aids. At first sight this may seem trivial. Yet there are countries where the slow construction of school buildings creates huge classes, where textbooks are prescribed but are not actually received by the teachers of the pupils, where tape recordings or language labs exist but have long since irretrievably broken down for lack of skilled technicians, where overhead projectors have been supplied but cannot be used because it is too costly to replace lamp bulbs, where even blackboards are almost illegible through old age and desuetude. Although such shortcomings are trivial they deeply affect the morale of the teachers, and through the teachers affect the learning prospects of the learners. But where the educational system is well-ordered, well provided and efficiently maintained, learning rates tend to be high—all else being equal.

6. *Realistic aims, accepted by all.* It is self-evident that the establishment of aims which are realistic should improve the climate of learning and teaching. But it is also necessary to state that these aims should be known to, and accepted by the teachers as a whole and enlightened public opinion in addition. Otherwise one is preparing the way for failures such as those mentioned in relation to some areas of primary school French in England.

7. *Suitable syllabuses.* The syllabus is potentially the most important single pedagogical formulation to bear upon any learning/teaching situation. We might define the ideal syllabus as a statement of *aims, approach, content, sequence* and *preferred methodology* for a given set of learners and teachers. With so many elements to be incorporated within a single formulation there are many opportunities for inadequacy. Yet our understanding of the principles of syllabus design has recently been revised, extended and modernised in ways that now make it possible to provide, for any course in learning and teaching languages, a formulation of great potential benefit. When this is done it can be a major source of guidance and support for the teacher, and through the teacher, a major influence upon the learner's progress and achievement. (This is an area where Britain has specialised in recent years. An example of the most sophisticated work is to be founded in the 4 volumes of *The Edinburgh Course in Applied Linguistics* (Oxford University Press, 1974-1977) where the ultimate value of applied linguistics, as far as TESOL is concerned, is shown as lying in its use for specifying the most suitable syllabus for a given set of learning/teaching requirements. The work of Pit Corder, John Trim, Henry Widdowson, John Munby, among others also contributes to this variable.) A suitable syllabus, in current terms, provides a head-start towards higher rates of learning achievement.

8. *Intensity of teaching relatively high.* Although virtually no research has been carried on the subject, there is widespread belief among teachers who have tried it that an increase in the intensity of teaching leads to a more than proportionate increase in the rate of learning per unit time. There seems to be a lower limit of around 5 or 6 hours per week, below which the learning yield per hour drops off drastically. The range of intensity from this 5-hour point up to another and rather different borderline, at around 20 hours per week, seems to constitute what one might call the normal range of intensity. 'Normal' here means that few alterations of method are required because of increases in intensity. above 20 hours per week there is some experience that special precautions need to be brought in, to guard against fatigue, boredom, too-long contact with the same teacher, and other technical difficulties. If the necessary precautions *are* taken, successful learning at intensities up to 40 or 50 hours per week, sometimes even beyond, have been reported. Within the 'normal' range of 5-20 hours per week it is widely accepted that the learning yield per hour increases directly with intensity: i.e. more learning takes place in 20 hours of instruction when it takes place in one single week than in 20 hours when they are spread over 4 weeks. Perhaps part of the reason for this is that the lower rate of intensity includes four weeks of opportunities for forgetting. The assertion being made here is that rates of achievement are improved by higher rates of intensity in teaching and learning.

9. *Teachers have a high level of professional competence.* This may again seem self-evident, except for those who deny the need for any intervention in the learning process. But what I am referring to is something quite specific, though difficult to describe. If we consider the full range of professional training as a language teacher it is beginning to look as if a quantum jump in effectiveness in the management of learning takes place at a particular point—a point which relatively few teachers have reached until recently but which is becoming increasingly common. In Britain, at least, this watershed point is generally associated with an opportunity for intensive *further* professional training following some 7 to 10 years of experience (e.g. typically in a one-year Master's course in Applied Linguistics or TEFL/TESL). Some rare, gifted teachers achieve this out of their own personal resources. At all events, it seems that a process of in-service professional maturation slowly takes place in the mind and understanding of the practising teacher, at the end of which the opportunity for some months of intensive thinking, reading and discussion produces a fresh burst of competence, confidence, and morale. This applies to teachers from any country, not just from Britain or the United States. I have no hesitation in maintaining that teachers who have achieved this professional regeneration become exceptionally well able to maximise the progress of their learners.

10. *Teachers cherish learners.* The verb *to cherish* appears very rarely in the literature of language learning and teaching. Its use here implies that the best teachers *know* their pupils, *encourage* them, *show concern* for them, *find*

out their interests, discover their learning preferences, monitor their progress with a sympathetic eye, *unravel their difficulties—cherish them* as human beings engaged in a collaboration of learning. There is a rough analogy here with 'Intensive Care' in hospitals, where the patient is constantly watched by skilled professional people whose first concern is to help the patient to want to live. *We* are concerned not only with helping our learners to learn, but with ensuring that even when they experience great difficulties they still *want* to learn.

Of all the strong recurrent reasons for success, this last is the most important and the most effective, and when it occurs it is frequently associated with the high level of professionalism referred to above, since it is normally teachers at this level who are not only best able to sense the learner's immediate needs but also have the broad range of professional skills to enable them to compensate in some degree for deficits elsewhere in the total learning/teaching situation.

The Methodological Status of the Conditions for Success

In this final section of the paper I wish to show that these inventories of important factors, and especially the phrase 'teachers should cherish learners,' are not simply pious assertions and exhortations to virtue, but that they have an intellectual status in relation to a model of the language learning/teaching process. Methodologically speaking (by that I mean in terms of ordering an argument, not in terms of teaching technique) there is a parallel between a phrase like 'success comes when teachers cherish learners' and an expression in, say, theoretical physics, such as $e = mc^2$. The methodological parallel is that in this phrase, as in the expression in physics, each term needs to be expanded into a complex statement, the various statements having complex inter-relationships, all of which have meaning in reference to a theoretical model.

So let us expand each term of the phrase 'success occurs when teachers cherish learners'.

(i) *Success comes....*	must be re-written as 'maximum rates of achievement in the learning and teaching of a foreign language are typically produced . . .'
(ii) *when teachers....*	re-write as 'when skilled and devoted teachers . . .'
(iii) *cherish....*	here we need to restore a deletion which bears upon the requirements of the social situation which alone makes these events possible: '. . . are encouraged by society and their profession to cherish . . .'
(iv) *learners....*	re-write as 'willing learners'.

The statement now reads, in its expanded form:

Maximum rates of achievement in the learning and teaching of a foreign *language are typically produced when skilled and devoted teachers are encouraged by society and their profession to cherish willing learners.*

But this is only the first stage of expansion of meaning. The statement still embodies many pre-suppositions which are made explicit in the theoretical model from which this statement derives.

I say 'derives' deliberately, for the following reason. The question of reasons for failure and conditions for success has interested me for over twenty years, but after trying in vain for half-a-dozen years to see a pattern in the complexity of language learning/teaching events I put the question aside and turned instead to a different issue, the elaboration of a model of the process. My reasoning was that if we could identify the absolute minimum elements in the process of learning and teaching languages, and could show the principal inter-relationships between them (i.e. if we could develop a model) then perhaps we would be better able to understand what is happening in language learning, and hence better able to answer the question about conditions for success. And so it turns out.

An Outline Model of the Language Learning/Teaching Process

In order to explain why the development of a model has made it possible to identify the most important factors—out of the whole diversity of LL/LT—for failure and success, I will briefly outline the nature of the model and its principal elements.

Figure 1 depicts a model of the process of learning a language with the mediation of a teacher: it is thus not concerned with first language acquisition. (This is not to deny any connection between acquisition of the mother tongue and learning a subsequent language: it is merely to say that mother tongue acquisition is not the focus of attention in constructing this model.) The process is seen as originating in two related elements: the first combines the sociolinguistic situation with administrative action—if you like, it brings together the public will that language should be taught with the provision for teachers to be trained and employed for that purpose—while the second element brings in the professional resources of language teaching: teachers of languages in a given country do not exist in a vacuum, they are part of a world-wide network of centres of knowledge, and they draw on the resources of linguistics, psychology, social theory, educational philosophy, etc. (In case anyone should ask what place linguistics has in the model, my reply is that it has an important place in this second element, as one among several crucial contributing disciplines.)

Next the process feeds a group of several elements, all concerned with different aspects of *teaching*. The first of this group of elements on teaching defines the minimum qualities of the teacher; the second deals with *approach*—the ideology, as it were, of language teaching; the third relates to *methodology* and to the whole gamut of instructional techniques; the fourth of them

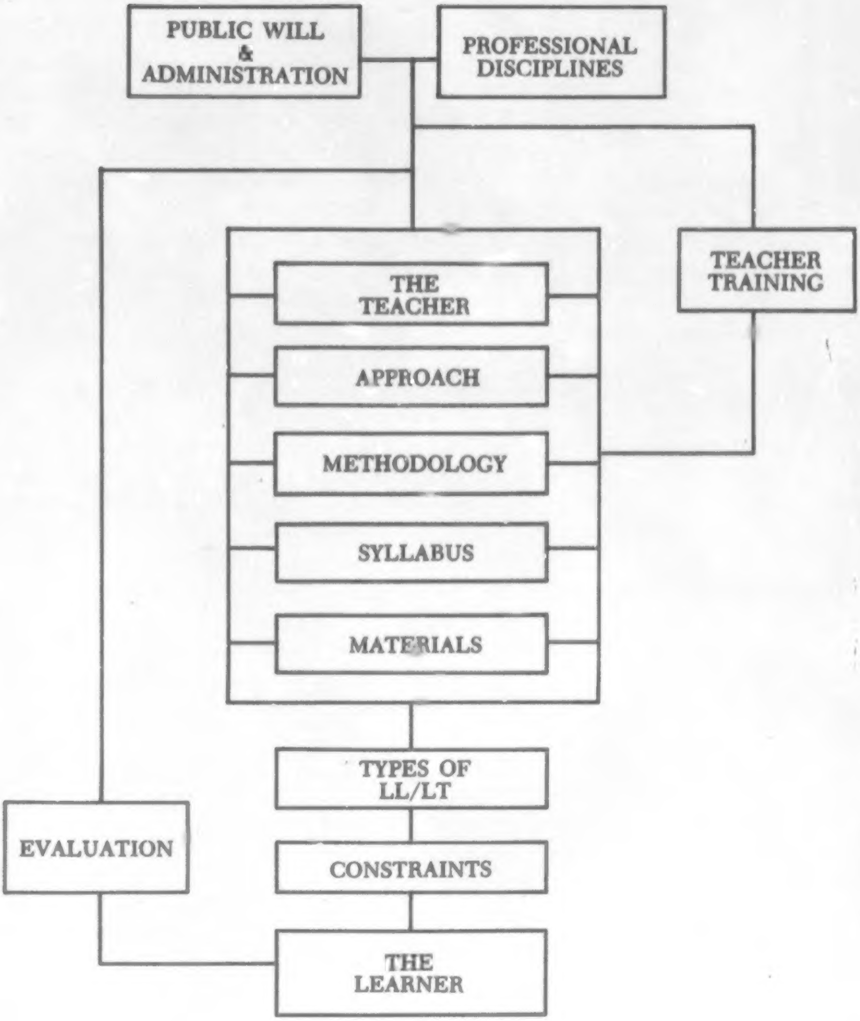

PUBLIC WILL & ADMINISTRATION
PROFESSIONAL DISCIPLINES
TEACHER TRAINING
THE TEACHER
APPROACH
METHODOLOGY
SYLLABUS
MATERIALS
TYPES OF LL/LT
EVALUATION
CONSTRAINTS
THE LEARNER

Figure 1. A Model of the Language Learning/Language Teaching (LL/LT) Process (adapted from Strevens, 1977).

relates to *syllabus;* the fifth to *materials;* and leading into all these is a seventh element defining *teacher training*. There is also an element dealing with *evaluation and assessment* as a feedback mechanism for the learning and teaching process.

Between the end of the block of elements concerned with teaching and arriving at the final element, the *learner*, the process goes through two variable filters or selectors. The first of these selects the appropriate type of LL/LT, according to parameters such as the age of the learner, the type of educational framework he is within, the level of proficiency he has reached, etc.; the second selecting element accounts for the various *constraints* and impediments that we looked at earlier—physical and organisational.

What reaches the learner, then, is the public will backed by administration, amplified by pedagogical expertise, selected for different types of learning and diminished by only those impediments, which are currently unavoidable.

And what of the *learner?* The final elements of the model deals with him (or her) as in a sense the focus of the entire process, and this element makes allowance not only for his possession of the universal potentiality for learning language but also for his individuality, his particular abilities, deficiencies, experience, personality.

Conclusion

The conception of failure and success in language learning which I have attempted to outline in this paper is based upon the notion of the whole and individual learner, not just upon the universal but common characteristic of acquisition; together with a model of the learning/teaching process. And if I were asked to abandon the whole of this paper and were allowed to retain one single sentence, it would be this sentence: *Maximum rates of achievement in the learning and teaching of a foreign language are typically produced when skilled and devoted teachers are encouraged by society and their profession to cherish willing learners.*

Teacher Feedback on Learner Error: Mapping Cognitions[1]

Michael H. Long

Concordia University

Knowledge of results appears to be an essential part of any human learning process, including language learning. Increasingly, the learner of a first or second language is believed to progress by the formulation of hypotheses about the target language grammar, based on the input data received, the use of feedback to test those hypotheses, and their subsequent rejection or modification. Not surprisingly, therefore, the provision of some sort of feedback (error detection and/or error correction) was one of only two characteristics found by Krashen and Seliger to be common to all language-teaching methods known to be successful. This paper reviews some recent studies of the ways in which learners in classroom settings are given information about the relative success or failure of their attempts to use the new language.

Options open to teachers when errors occur are described in the second part of the paper, and a model is presented of the decision-making process. Several factors bearing on these decisions are considered. The model shows, among other things, how different decisions can sometimes result in overtly identical teaching behaviours, thereby presenting students, teacher-trainers and researchers into classroom processes with some tricky problems of interpretation.

Language learning as a process of hypothesis-testing and rule-modification is reconsidered in the light of what appears to be a marked lack of clarity and consistency in teacher feedback. The paper questions the status of "correction" of error as an essential characteristic of successful classroom language instruction.

The purpose of this paper is briefly to consider the importance of the role of feedback in any cognitive theory of second language learning, and then to review some recent descriptive studies of the classroom behaviours of teachers following learner error, in an attempt to ascertain what the teachers studied currently do when providing feedback. I will then set out what I see as the basic options open to teachers when giving feedback on error and describe some of the factors affecting decisions when choosing among the options presented. In the absence of any studies of the effect of different forms of feedback on student learning, I will, clearly, be unable to recommend one decision or series of decisions over another; I will, however, question the status commonly attributed to so-called teacher "correction" of error as an essential characteristic of successful classroom foreign/second language instruction.

---

[1] At various stages in the preparation of this paper I received valuable comments from Richard Allwright, Fernando Castaños, Craig Chaudron, Alison d'Anglejan, Patsy Lightbown and Brian Smith. The opinions expressed do not necessarily reflect theirs, and all errors are, of course, my own.

The Role of Feedback

In order to learn how to perform even the simplest task people require information on the success or failure of their attempts at performance. The form of that feedback can vary greatly. Thus, an infant may learn to coordinate the movements needed to stretch out its arm in order to pick up an object by seeing its hand touch the object and by feeling and perhaps hearing the contact made between object and hand. The function of knowledge of results (KR), too, can vary. Annett (1969) has observed that its effect on the learner may be primarily that it provides him or her with (a) information and/or (b) reinforcement, and/or (c) some form of incentive.

As language teachers we give tacit recognition to the importance of KR both by the amount of time we spend informing our students of when, where and (sometimes) why they have gone wrong, and by the wide range of techniques we have developed for doing so. Krashen and Seliger (1975) found that the possibility of some sort of feedback, error detection (by the learner) or correction (by the teacher), was one of only two characteristics common to all language-teaching methods known to be successful. By implication, feedback on error[2] by the teacher or some teacher surrogate is potentially one of the necessary and sufficient conditions for successful foreign/second language learning. This and the universality of the practice makes it a subject worthy of careful study; one would like to know whether we achieve by it what we *think* we do and if there are any ways of making the process more efficient.

We are interested in error detection, i.e. KR, rather than the narrower error correction alone, for we are all familiar with the ability of some of our students obstinately to repeat the wrong answer despite our dogged provision of the correct one. It would boost our confidence to know that, even if a learner could not get an utterance right as a result of our provision of feedback, he at least knew when he was wrong. For present purposes error 'correction' is too limited a term for other reasons, describing as it does the (hoped for) *result* of feedback on error, not the feedback itself. This underlines the importance of viewing feedback in its proper context, that is, following error and prior to further attempts at correct student performance. What is sometimes described in the literature as "a correcting move" should be seen for what it is: behaviour by the teacher which allows the learner to obtain KR, on the basis of which, hopefully, it will be the *learner* who makes a correcting move. Feedback is designed to promote correction but is not itself correction. Correction occurs, according to the hypothesis-testing model of language learning (see below), when the learner modifies a rule in his or her interim grammar of the language being learned.

---

[2] In this paper an 'error' will refer to (1) any phonological, morphological, syntactic or lexical deviance in the form of what students say from a standard variety of English which is attributable to the application by the learner of incorrect grammatical rules, (2) recognisable misconstrual of or lack of factual information, (3) a breach of rules of classroom discourse, and (4) a bit of student language behaviour treated as an example of (1), (2) or (3) by the teacher.

The study of feedback on error is of theoretical as well as practical interest. Increasingly, the learner of a first or second language is believed to progress by the formulation of hypotheses about the target language grammar, based on the input data received, the use of feedback to test those hypotheses, and their subsequent rejection or modification. Language learning has not always been viewed in these terms.

Due largely to the predominant influence of neo-behaviourist theories of learning on foreign language education in many (but not all) parts of the world during the 1950's and '60's, the importance attributed to KR was for some time limited to its perceived role in reinforcing emergent (correct) second language habits. However, with the re-evaluation of the language-learner's intellect by Corder, Nemser, Dulay and Burt, Selinker, Richards, Taylor and others, the role of errors and of feedback on error has taken on a new significance. Of errors, for example, Corder wrote in his classic 1967 paper:

> Errors are indispensable to the learner himself, because we can regard the making of errors as a device the learner uses in order to learn. It is a way the learner has of testing his hypotheses about the nature of the language he is learning. (Corder, 1967, p. 167.)

And of feedback, eight years later:

> In order to test this hypothesis of his—this is in no sense, of course, a conscious process—he makes utterances which are generated by his particular interlanguage grammar at a particular moment. *The behaviour of the teacher or other speakers of the target language enables him to decide whether any particular hypothesis he has developed is valid or not.* (Corder, 1975, p. 411. Emphasis added.)

At first sight, far from undervaluing the learner's potential cognitive contribution to the learning process, the above statement would seem to attribute to him or her part of the ability of the professional linguist. Corder and others are careful to emphasise, however, that the hypothesis-testing is in no sense a conscious process. Nonetheless, it is clear that the way in which language teachers behave following their students' commission of error is potentially of interest for any cognitive theory of the second-language-learning process aided by formal instruction.

Teacher Feedback on Learner Error

Several writers on child language development in the sixties reported the focus of mothers on the communicative effectiveness rather than on the formal accuracy of what their children said. (See, e.g. Brown, Cazden and Bellugi, 1969.) They were observed to expand child utterances, adding missing functors, but seemed to do so to confirm that communication had taken place rather than to provide specific language instruction. Some more recent studies, (e.g. Nelson, Carskaddon and Bonvillian, 1973; Moerk, 1976), suggest that deliberate "teaching activities" can play a role in first language acquisition. In spontaneous mother-child interactions in the home, Moerk observed such activities as mothers modelling little question and answer sequences for the child, and what Moerk calls 'prodding', where the mother urges the child to make some improvement with "Say X" or "Can you say X?", where X is the model of the construction she wants the child to imitate. These data appear to argue against an outright dismissal of negative feedback on formal inaccuracies as a factor in successful Ll acquisition—and, by implication, in L2 learning with formal instruction. However, even with Moerk's data one is struck by the presence of a message-clarification element often lacking in second-language classrooms. Further, what appear to be "language-teaching" sequences may not have this function but again that of establishing (lack of) communication. Also, showing that mothers do this kind of "teaching" does not prove that children learn by it, unless, that is, one is prepared to accept apparent improvement of some kind in the child's next utterance as evidence of learning. This would seem unjustified given what longitudinal Ll acquisition studies show about the gradual emergence of native-like forms, errors co-occuring with correct production over periods of several months. (See, e.g. Slobin, 1971.)

In the last few years there have been a number of participant observational studies of teacher behaviour following errors by second language learners. They have mostly described the feedback practices of experienced teachers working with students of differing ages and levels of language proficiency, often in relatively small groups, and with English or French as the second language being taught. Some of the findings of these studies are reported and discussed below. Descriptive studies only are considered here as it would seem reasonable to suppose that the value of the provision of feedback on error, like any other teaching behaviour, will only be measurable once we have identified what teachers actually do in classrooms, and can distinguish this from what they are sometimes imagined to do or are urged to do by textbook writers, teacher-trainers or the authors of methods books.

Teacher reacting moves following learner error have been observed by several investigators to lack clarity. In a pioneering study of teacher feedback on error in oral work, Fanselow (in press) videotaped eleven experienced teachers giving the same lesson on adjective order, ("He's wearing a grey plastic raincoat," etc.). Verbal and non-verbal behaviours were analysed. It was found that, following student errors, teachers often gave more than one form of feedback simultaneously, e.g. gesturing "no", rubbing one item of clothing (to indicate the material from which it was made) and pointing to another object to be described, or saying "fine" while shaking the head sideways. As Fanselow points out, even if students *were* able to interpret the intended message correctly, (factually correct answer but with one or more errors of form), they would be no wiser as to what was wrong with their utterance. Subtle shakings of the head and "Again" might have meant "I didn't hear you—please repeat" or "You made a mistake". Such messages were often given while students were looking at fellow students, at the learner who had just committed the error, at materials, or anywhere but at the teacher. Thus, they can have meant nothing to the learner for being unseen.

Fanselow also found clarity to suffer due to the teachers' frequent use of the same overt behaviour for two or more purposes, a phenomenon noted by several investigators. Typical is the provision by the teacher of the model response desired immediately after an unsuccessful student attempt at production. Fanselow offers the following example:

A. 1 T : It's blue. (SAMPLE)
2 Sl : It blue.
3 T : It's blue. (MODEL FOR COMPARISON)
4 S2 : It's blue.
5 T : It's blue. (CONFIRMATION)
6 Sl : It blue.
7 T : It's blue. (MODEL FOR COMPARISON)
8 Sl : It blue.

The same teacher utterance sometimes serves as (1) a sample of the target language data to be imitated by the student, (line 1), as (2) a model with which the learner is supposed to compare his own imperfect response, and, by implication, an indication that an error has been made, (lines 3 and 7), and as (3) confirmation and (some people would say) "reinforcement" of a correct response, (line 5). In fact, as Allwright (1976) has noted, function (1) characterises everything a teacher says in the target language, 'guidance' as to its nature simultaneously constituting a further language sample. Teachers sometimes attempt to distinguish overtly identical but multi-functional utterances of the 'It's blue' variety by changes in intonation, e.g. the use of a rising "question" tone or the addition of emphasis for case (2) above, (model for comparison), and/or by accompanying non-verbal cues. From the point of view of lack of clarity, however, a learner's inadequate grasp of English intonation may often prevent him from making the same (or any) distinction. Further, even if interpretation (2), i.e. "Wrong—listen to what you should have said", is correctly understood, the learner may again be unclear as to which part of what he said was wrong. Neither the source, nor location nor identity of error has been indicated—simply its presence.

Here is another example[3], this time from an elementary level ESL class for ten-year-olds during practice of adverbs of frequency:

B. 1 T : All together.
2 T & SS : Yes, I always use a toothbrush to brush my teeth.
3 T : OK. David, can you repeat?

---

[3] Extracts from transcripts not attributed to other sources are from preliminary data gathered by the author and Brian Smith of the TESL Centre, Concordia University.

4 Sl : It's brush eh–
5 T : Yes, I always use a toothbrush to brush my teeth.
6 Sl : It's always toothbrush and it's my teeth.
7 T : I always use a toothbrush to brush my teeth. Marie?

David's possible belief that his second attempt (line 6) was correct will have been strengthened if he interpreted either his teacher's third statement of the correct form (line 7) or the transfer of attention to another student as confirmation/"reinforcement". Even if he interpreted the feedback correctly, the obvious disparity between what he is currently capable of and what he is being asked to produce makes it extremely unlikely that he was able (a) to remember what he said, (b) to do so long enough to compare it with what his teacher said, in order (c) to spot the differences and (d) to modify his future attempts at production. (See Allwright, 1975a, for a discussion of these and related problems.) That is, the teacher's feedback may not have been very informative, even if interpreted by the student as having this purpose at all.

Another often reported source of ambiguity concerns teachers' use of positive feedback, usually in the form of praise markers. The next example, concerning feedback on a pronunciation error, is from a corpus of data from four classes at the University of Essex, where experienced ESL teachers were working with adult Venezuelan students. It is quoted by Stokes (1975):

C. 1 T : Again.
2 SS : The fifth of January, nineteen seventy-four.
3 T : I think, Eulyces, I heard something else here. Will you say it alone?
The . . .
4 S2 : The . . .
5 T : The first.
6 S2 : No, the [fɪf]
7 T : Fifth?
8 S2 : [fɪf]
9 T : It's very hard to say . . . fifth.
10 S2 : The [fɪs] of January, nineteen seventy-four.
11 T : Good.

Notice the redundancy (at least in terms of feedback on the error the teacher supposes s/he has heard) of lines 3 to 5. The teacher's communication of the existence of error (line 3) is, in effect, an invitation for the student to commit it again, and is lacking in information as to its location or identity. Of the teacher's use of 'Good' in line 11, Stokes comments:

> One suspects the approval in line 11 is a measure of desperation on the part of the teacher. However, the word was used on a number of occasions when the response was incorrect. One might suggest that "Good" at times simply signalled a boundary marker and did not carry any meaning of commendation. But how does

> the student interpret the word? If he knows that his answer was really not correct, is his confidence in his teacher weakened? If he believes that "Good" indicates a correct response, will the student continue with the error? (Stokes, 1975, p. 61.)

Loftus writes of the use of the "praise markers", "Good", "Fine" and "OK" in more of the Essex data:

> The word "praise" is really a misnomer for these words, as very often it is really a question of encouragement, conciliation, or simply use as a boundary marker. (Loftus, 1975, p. 81.)

In our own data, the same ambiguity occurs with "Yes" and "Mm hm", which are sometimes used to interrupt a student, (see, for example, extract B, line 5), sometimes to confirm correctness and sometimes to indicate the existence of error, then usually accompanied by changes in intonation, voice quality, gestures, facial expressions or combinations of these.

Another source of lack of clarity is, of course, the sheer linguistic or conceptual complexity of some of the feedback teachers give. This has already been alluded to in the previous discussion of the teacher reacting moves in extract B, lines 5 and 7, and extract C, line 3. Most of us are familiar, too, with "explanations" of points of grammar that can be conceptually taxing even for a linguistically unsophisticated native speaker. Both teachers and textbook writers occasionally add to the student's burden by couching the ideas in metalanguage more complex than the point of grammar under discussion.

The lack of clarity of individual feedback moves is often compounded by the inconsistency in a series of such moves. (It may be, of course, that the inconsistency is well-motivated, as we shall see later.) In some of the lessons studied, *what* is treated as erroneous, *who* is so treated and *how* that message is transmitted sometimes appear, superficially at least, somewhat arbitrary. Fanselow (in press) found teachers focusing students' attention on the omission of articles or auxiliaries in one part of a lesson but ignoring the same errors in other parts. Stokes (1975) cites this example concerning the definite article:

D. 1 S3 : When did you leave Venezuela?
2 S2 : I left Venezuela eh eleventh of January.
3 T : Good.

followed later in the same lesson by:

4 T : When was he born?
5 S2 : Twenty . . . twenty-first of January, nineteen sixty-three.
6 T : Come on, Eulyces, you missed something here. Just say it over again.
7 S2 : Twenty . . .
8 T : The twenty-first.
9 S2 : Twenty-first of February, nineteen sixty-three.
10 T : Good.

Stokes wonders what hypotheses the student can be forming. The first article omission was praised (lines 2 and 3), the second reprimanded (lines 5 and 6), and the third approved again (lines 9 and 10).

The same inconsistency as to what is treated as error and also as to who is treated as having made an error was found by Mehan (1974) in his study of two "orientation lessons" for early primary-school-age children. Mehan reports the teacher's insistence on complete sentences for children's responses on some occasions, her acceptance of incomplete sentences on others, even from the same children, and her rejection of the complete correct response on others.

The following exchange occurred in an ESL class for ten and eleven-year-olds.

E. 1 T : What are you looking at? . . . The T.V.
2 S1 : I'm looking at oranges flowers.
3 T : OK. I'm looking at oranges flowers.
4 S1 : I'm looking at oranges flowers.
5 T : What is he looking at? Yes eh Peter?
6 S2 : He's looking at oranges flowers.
7 T : Flowers. Orange flowers. OK. David.
8 S3 : He's looking at orange he's looking at orange flowers.
9 T : OK. Good. Marie.
10 S4 : He's eh looking at eh oranges flowers.
11 T : Orange flowers.
12 S4 : Orange flowers.
13 T : OK. Orange flowers.
14 S4 : Orange flowers.
15 T : OK. All together. Orange flowers.
16 SS : Orange flowers.
17 T : He's looking at orange flowers.
18 SS : He's looking at orange flowers.
19 T : OK. Very good. You can go back to your place.

The whole sequence lasted 58 seconds. Student one's erroneous 'oranges flowers' is accepted (line 3) and repeated by the teacher, a non-native speaker, who temporarily adopts the wrong form herself, followed (in line 4) by a second production of the erroneous version by student one. This, too, is tacitly approved (line 5) by the teacher's transfer of attention to another student, Peter, who dutifully proceeds to perform the desired 'I'm→He's' transformation and copy the rest of the model sentence (line 6) complete with error. Whether student two, the teacher or anyone else in the class notices the difference between Peter's answer and the teacher's echoed acceptance (line 7) is unclear. The teacher has now reverted to the correct form and (line 8) student three proceeds to get it right, too. Marie, student four, does not. Her 'oranges flowers' is the same as two efforts by student one and one by student two which have been accepted by the teacher as well as modelled by her once; however, perhaps because of her hesitant rendition, it is now rejected by the teacher (line 11). Marie is obliged to repeat the correct version twice, alone, before treatment is transferred to the whole class.

It is interesting to note that, prior to Marie's first attempt at production, there had been four incorrect models (lines 2,3,4 and 6) and two and a half right ones (lines 7 and 8). A linguist working with the same data from an unknown language would, perhaps like Marie, have hypothesised that 'orange' and 'oranges' were free variants or that the pluralised adjective was the correct form because it was the more frequent one. Of course, extract E is (perhaps) atypical because of the teacher's momentary lapse. Nonetheless, the presence of grammatical errors in the classroom learner's input data, due to the presence in this of his fellow students' interlingual speech, is the rule rather than the exception, and has been noted (e.g. by Allwright, 1975a, and in some studies of immersion programmes, e.g. Plann, 1977) occasionally to be adopted in preference to the teacher's error-free production. In the same article, Allwright has discussed the multitude of conflicting hypotheses a learner may construe when erroneous forms are accepted from some students but not from others. Similar cases have been noted by Stokes (1975) and Mehan (1974). Loftus (1975) found differential use of the praise markers, 'OK', 'Good' and 'Fine', in accordance with his (the teacher's) perceptions of students' abilities and personalities.

The variety of ways in which feedback is transmitted is perhaps less likely to cause confusion than some of the other factors noted. Zahorik's study (Zahorik, 1968) of positive and negative feedback by teachers in 15 "content" subject classrooms revealed 175 different types, but only 16 of these occurred with any frequency. Also, the kind of feedback given was not found to vary randomly but with grade level, lesson purpose, the kind of learning tasks set and the quality of student response. The kinds of options taken up by second language teachers—remodelling the correct response, with or without emphasis of some elements, resetting the task, setting a simpler one, and so on—are reflected in the various category systems and fragments thereof, found in Fanselow (1977), Allwright (1975a), Cathcart and Olsen (1976) and others. Such systems do not show, however, that individual teachers may actually use a fairly narrow range of the options listed in category form in instruments purporting to provide exhaustive coverage of all teacher feedback behaviours observed during their creation.

Fanselow (in press) found that teachers entirely avoided some possible categories of feedback and that other forms categorised by him were employed by just a few of the teachers he observed. There was variation, however; only one form, presenting the right answer or part of it after an error, was found to be common to all the teachers studied. This coincides with Cathcart and Olsen's finding, in their survey of 21 teachers of ESL in California, that (re)modelling the correct responses was the most used option, (Cathcart and Olsen, 1976). The *consistency* shown here needs to be interpreted cautiously, however, in view of the findings concerning the multi-faceted use of this and other overt teaching behaviours. Also, in many ESL programmes, students move from one teachers to another with a change of course and/or receive instruction from more than one teacher within a given course. They have to learn how and when each teacher gives feedback and to interpret accordingly. Nevertheless, there is some evidence that groups of teachers, however vaguely defined, may share characteristics. In her study of ten female ESL teachers in Israeli secondary schools, Lucas (1975) found variations in the quantity and type of errors treated. Native speakers of English treated more lexical errors than any other kind; non-native speakers treated more pronunciation errors. In general, native speakers tended to ignore more, to treat less and to disapprove less than did the non-native speakers. In a study of a total of twelve half-hour lessons in grades eight and nine by three teachers in French immersion programmes in Canada, Chaudron (1977b) also found consistency in the (greater) relative importance attached by teachers, as reflected in their feedback moves, to the learning of lesson content (science, math, geography and history) over errors in the students' French, except in French classes, where there was relatively more equal treatment of grammar and content errors. In this respect, the content subject teachers were behaving more like mothers interacting with children acquiring their Ll.

The data presented need to be viewed in perspective. Both the numbers of teachers studied and the periods of observation have been small, and so the extracts cited are not necessarily representative samples of most teachers' feedback practices or even of *these* teachers' normal, i.e. unobserved styles. Such generalisations as have been made, must be interpreted cautiously and should be understood only as testable hypotheses about classroom language teaching in general, not as facts about what teachers everywhere do. Further, as discussed in the section below, much of the apparent lack of clarity and consistency may be just that—apparent—and attributable to sound pedagogical practice.

Options and Decisions

Short stretches of recorded verbal classroom interaction like those examined lose much by being taken out of the context not only of the lessons from which they come but of the whole history of the relationships established prior to those lessons between the teacher and students concerned. When an error occurs and is noticed by the teacher, he or she is faced with several complex decisions the making of which require the careful weighing of many factors often not apparent in the transcripts even of complete lessons. Just some of these are the teacher's awareness of his or her usual (previous) feedback practices with these students, and beliefs as to the students' familiarity with them, his/her beliefs as to their success so far (particularly with the student who has just committed an error) and perception of the measures' popularity among the students in the class, (including, again, the error-making student). Other factors considered include the objectives of this lesson and the course as a whole, the pedagogic focus at the moment of error commission and the teacher's perception of the likely outcome of treatment in terms of the error-making student's aptitude, personality, ability and sociometric status.

To take just one of these considerations as an example, it is a rare teacher indeed who is blessed with a class of students of equal proficiency in the target language and equal ability for learning more of it. Mixed ability groups are taken into account (among a lot of other ways) by the distribution among the students of questions of different complexity and by the acceptance of different standards of performance from them. Student X rarely seems to get anything right and, so that he does not give up altogether, he is set simpler performance tasks and given encouragement for almost anything he produces. Student Y, on the other hand, has a superior command of English, picks up new pieces of it as fast as they are presented to him and can easily become bored if not set demanding learning tasks. His aims and the teacher's for him are higher; thus, when he commits an error—any error—amid an otherwise near-perfect performance, he is informed of the fact in the knowledge that he will (and will want to) attain a still higher standard of accuracy. This example of well-intentioned differential treatment, (resulting from teacher perceptions of just one kind of individual differences among learners), is one of many potential explanations of the inconsistencies in the teacher feedback in, for example, extracts C, D and E.

Figure 1 is a model of the major steps in the decision-making process teachers go through prior to performing some overt (and so directly observable) behaviour following learner error. The remainder of this section is devoted to a discussion of these steps and an attempt to show how various factors can influence the making of the decisions.

Teachers do not provide feedback on all errors that are committed. For a variety of reasons many errors go unnoticed; it is only those perceived by the teacher which serve as input to the decision-making process. Having noticed an error, the first (and, I would argue, crucial) decision the teacher makes is whether or not to treat it at all. In order to make the decision the teacher may have recourse to factors with immediate, temporary bearing, such as the importance of the error to the current pedagogic focus of the lesson, the teacher's perception of the chance of eliciting correct performance from the student if negative feedback is given, and so on. Consideration of these ephemeral factors may be preempted, however, by the teacher's beliefs (conscious or unconscious) as to what a language is and how a new one is learned. These beliefs may have been formed years before the lession in question.

If the teacher decides to ignore the error and, for example, to set the same or a new learning task, notice that the overt teaching behaviour can be the same as that if he or she failed to notice the error at all or noticed but decided to postpone treatment until a later lesson or part of the current lesson. An observer will not be able to tell the difference, except indirectly through a subsequent interview with the teacher. Note, too, how teachers favouring broadly inductive or deductive teaching strategies could weigh such factors as pessimism

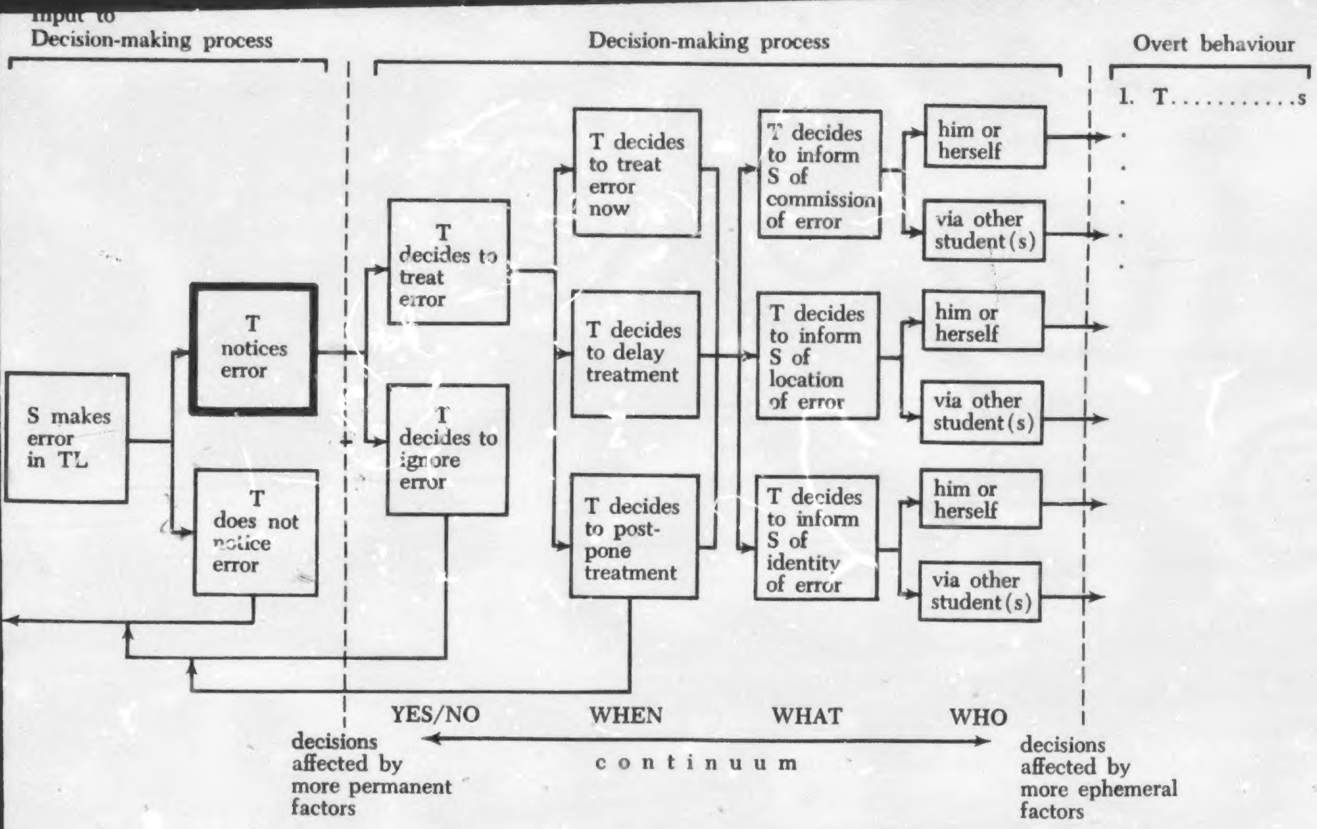

Input to
Decision-making process
Decision-making process
Overt behaviour
1. T..........s
S makes error in TL
T notices error
T does not notice error
T decides to treat error
T decides to ignore error
T decides to treat error now
T decides to delay treatment
T decides to post-pone treatment
T decides to inform S of commission of error
T decides to inform S of location of error
T decides to inform S of identity of error
him or herself
via other student(s)
him or herself
via other student(s)
him or herself
via other student(s)
YES/NO
WHEN
WHAT
WHO
continuum
decisions affected by more permanent factors
decisions affected by more ephemeral factors

Figure 1. Model of the decision-making process prior to the teacher feedback move.

about the likelihood of improved student performance over the importance they attach to their general approaches to teaching languages, and opt for the same reacting move, e.g. resetting the same task for another student. That is, the same overt teaching behaviour could be appropriate at a given moment for either, despite the different quantity or quality of feedback one might expect under either basic approach.

Assuming the teacher decides to treat the error in some way, the next decision concerns *when*. Here the options appear to be (a) immediately, which often involves interrupting the student, (b) after the (apparent) completion of the student's utterance, or (c) at some future time, including, for example, in the course of a future lesson. Chaudron (1977a) calls (b) and (c) 'delayed' and 'postponed' feedback, respectively. There is some evidence in the psychology literature that the value of feedback decreases the greater the time lapse between performance and KR, and that KR given verbally by an experimenter after each response or failure to respond is effective in preventing the normally observed decline in vigilance over time, (Mackworth, 1950). There is, however, the important question of the inhibiting effect of interruptions, and their inevitable communication to the student of the greater value placed by the teacher on the form as opposed to the content of what he or she says.

In favour of choosing to delay or postpone treatment, there is also some evidence of the positive effects of what is known as "wait time". Hoetker (1968) found that in classes of above average ability English teachers allowed more time for students to begin their responses to soliciting moves than they gave students in below average classes. Rowe (1969), studying science lessons, reported that if average wait time was prolonged to five seconds or more, the length of student responses increased, as did the number of complete sentences produced, the confidence of those utterances as judged by their tone, the amount of speculative thinking verbalised by the children, and the number of children who engaged in that kind of talk.

In the second language classroom, Holley and King (1971) have described some innovations they made in the provision of feedback in the teaching of German to graduate students. Feedback was given on students' success in communicating 'content' in the foreign language rather than on the grammatical accuracy of the forms used to do so. Specifically, teachers were told not to interrupt students trying to answer questions, and when they hesitated in answering, to say and do nothing (except wait) for from five to ten seconds. Only then were they to (a) rephrase the question, reducing where possible the number of words while consciously emphasising content words over functors, or (b) cue the correct answer using grammatical variations of a key content word, or (c) encourage other students in the class to 'generate' simple sentences from the faltering students' incomplete utterances. They were also urged to expand ungrammatical but meaningful answers in the manner observed in parents interacting with young children, but not to demand repetitions of such expansions. Holley and King report that in over 50% of the instances they filmed, no "corrective measure" (rephrasing, cueing, 'sentence generating') was needed. Simply allowing students sufficient time to reformulate their responses led to improved performance.

Having made the decision as to *when* to treat errors the teacher then has three basic options, (a) to inform the student of the existence of error, (b) to inform the student of the location of error, or (c) to inform the student of the identity of error. Choice (b) will entail (a); (c) will entail (a) and (b). Examples of overt behaviours realising these options might be:

S : He go to the park on Saturdays.
(a) T : No.
or T : He go to the park on Saturdays? (I.e. the student's utterance repeated with rising intonation, probably accompanied by some non-verbal cue, such as raising eyebrows.)
(b) T : He *go* to the park on Saturdays?
or T : He *what* to the park?
(c) T : 'Go' or 'goes'?
or T : You missed the third person 's' off 'goes'.

There then follows a decision as to whether (a), (b) or (c) will be carried out by the teacher or by another student or students (column (4)—'WHO'—in the figure). There is no obviously recommendable path through columns (3) or (4), even for specific types of learners. Given the abundant possibilities for lack of clarity of feedback, however, other things being equal, there may be a case for more attention to be paid to option (c), with, where appropriate, an explanation of the source of error. Evidence as to the desirability of peer feedback is attractive, but this is but one of a variety of considerations affecting choice among the options in column (4). It might be outweighed, for example, by knowledge that other students are making the same error, or that the student concerned is unlikely/unwilling, e.g. for cultural reasons, to listen to his peers in a case like this.

Allwright (1976) has presented an 'inductive model incorporating cognitions' of the conditions necessary for successful learner production of the *response* part of the (teacher) Initiation—(student) Response–(teacher) Feedback sequence in classroom discourse. He showed how, following a teacher soliciting move and prior to the student response, (two overt behaviours), the learner must be attending, must believe he has heard the teacher's "question" correctly, that he remembers and understands the soliciting move, and so on. All of these factors are covert, unobservable, the result of a logical analysis of what the learning situation described must demand of any learner in cognitive terms. In the second part of this paper I have followed a similar approach in suggesting what I believe to be the basic decisions a teacher takes prior to the observable act of providing feedback on learner error, and have indicated a few of the factors bearing on those decisions. I have not argued the case for any particular decisions or series of decisions for the simple reason that, like so much of what we do as teachers in classrooms, there is little or no empirical evidence relating some form(s) of feedback with student learning.

Some studies *have* talked of the success of various forms of feedback in terms of the extent to which, following them, the learner is observed to incorporate some immediate improvement into his next utterance. This seems a questionable yardstick by which to judge efficacy. It does not fit with what, as teachers, we know to happen in classrooms. Students' errors do not by and large change in such a decisive manner from one moment to the next. Incorrect forms disappear gradually, and then often reemerge later, much as they do in first language development; they are not there one moment and gone the next. Temporary correct (improved) performance of what Krashen, et al (1977) call "easy" grammatical items, e.g. third person singular 's', is understandable following feedback because such items are easily monitorable. At the second attempt, having got his message across the first time he spoke, the learner is concentrating (and specifically being asked by the teacher to concentrate) on formal accuracy, so improvement should be expected. But what of "hard" areas of grammar, such as article usage? A missing or wrongly chosen article is easily incorporated in an adult learner's speech when he is focusing on formal accuracy, immediately after feedback. It seems unreasonable to suppose, however, that the learner has mastered this complex syntactic domain due to his attention having been drawn (explicitly or implicitly in the teacher's feedback move) to one more instance of correct usage after so many previous examples have failed. There seems to exist a danger of confusing the efficacy of various forms of feedback in terms of their effect on monitored linguistic performance with their usefulness in terms of bringing about a lasting modification of the learner's interim hypotheses about rules in the target language grammar.

Conclusion

Intuitively, the hypothesis-testing model makes sense in terms of adult learners seeking positive confirmation of hypotheses about the target language grammar, and this interpretation is supported by work on simplification and generalization strategies by Taylor, Richards and others. Richards (1974), on the other hand, has pointed out that taking the model to suggest that learners seek out negative instances in order to test (in the sense of *falsify*) their hypotheses is counter-intuitive. People like to discover they are right about things, not mistaken. Setting out to elicit negative feedback in this way would be an emotionally unrewarding task, one that was of low productivity in communicative terms and a very inefficient manner of achieving anything of use. (As Richards observes, learners could, instead, simply ask whether a given sentence was not in fact wrong). Given the unclear, inconsistent and complex nature of much teacher feedback on error, as illustrated in this paper, it would also be unlikely to work. Yet the falsification idea seems to be the interpretation one would have to hold in order to suggest, as one or two writers have done recently, that teachers should encourage learners to make errors.

The role of teacher feedback, whether errors are in any sense committed intentionally or not, is clearly vital, but it may be that not all forms of feedback are as vital as we sometimes think. We know that, like children, adults can and do acquire second languages outside classrooms simply by exposure to their use in situations involving verbal interaction for communicative purposes, i.e. with the aid of little or no formal "teaching" of any kind, including the "correction" of ungrammatical forms. What adult acquirers do undoubtedly have access to is information as to the success or otherwise in communicative terms of their attempts at using the new language; that is, they are able to find out whether or not they are succeeding in transmitting or understanding a message and, on the basis of this knowledge (in some as yet unknown way) to work out where improvement is needed. I would suggest that the case for what Krashen and Seliger call error correction (as opposed to detection) as one of the necessary conditions for successful second language learning is unproven and will remain so until longitudinal studies are undertaken of the second language acquisition process aided by formal instruction with differing feedback conditions.

Elsewhere (Long, 1976), I have drawn attention to the deleterious effects on the naturalness of classroom discourse of, among other practices, the so-called "correction" of errors. The greater *efficiency* (if any) of language instruction which includes various kinds of feedback on formal accuracy needs to be tested, but meanwhile can probably be improved by teachers reconsidering the options available to them and the decisions they make following learner error. Greater clarity and consistency of feedback, at least, must be desirable, if only on the grounds that if a job is worth doing, it is worth doing properly. Hopefully, further research will show whether the job is really worth doing at all.

REFERENCES

Allwright, R. L. 1975a. Problems in the study of teachers' treatment of learner error. In M. Burt and H. Dulay (eds.), *New directions in second language learning, teaching and bilingual education: On TESOL '75*. Washington, D.C., TESOL.

Allwright, R. L. (ed.). 1975b. *Working papers: language teaching classroom research*. University of Essex, Department of Language and Linguistics, September, 1975.

Allwright, R. L. 1976. Putting cognitions on the map: an attempt to model the role of cognitions in language learning. Paper presented at the Fourth AILA Congress, Stuttgart, August, 1975. In *Workpapers in TESL* 9, June, 1976, University of California, Los Angeles.

Annett, J. 1969. *Feedback and human behavior*. Harmondsworth, Penguin.

Brown, R., C. Cazden, and U. Bellugi. 1969. The child's grammar from I to III. In J. P. Hill (ed.), *Minnesota Symposium on Child Psychology*.

Cathcart, R. and J. Olsen. 1976. Teachers' and students' preferences for correction of classroom conversation errors. In J. Fanselow and R. Crymes (eds.), *On TESOL '76*, Washington, D.C., TESOL.

Chaudron, C. 1977a. The context and method of oral error correction in second language learning. A qualifying research paper prepared in partial fulfillment of requirements for entry into a Ph.D. program, O.I.S.E., Toronto.

Chaudron, C. 1977b. Teacher priorities in correcting learning errors in French immersion classes. *Working Papers on Bilingualism*, No. 12.

Chaudron, C. 1977. A descriptive model of discourse in the corrective treatment of learners' errors. *Language Learning* 27, 1, 29-46.

Corder, S. P. 1967. The significance of learner's errors. *IRAL 5*, 4, 161–170.

Corder, S. P. 1975. The language of second language learners: the broader issues. Modern Language Journal 59, 8, 409–413.

Fanselow, J. 1977. Beyond *Rashomon*—conceptualizing the teaching act. *TESOL Quarterly 11*, 1, 17–39.

Fanselow, J. 1977. The treatment of error in oral work. *Foreign Language Annals 10*, 4, forthcoming.

Holley, F., and J. King. 1971. Imitation and correction in foreign language learning. *Modern Language Journal 55*, 8.

Krashen, S., and H. Seliger. 1975. The essential contributions of formal instruction in adult second language learning. *TESOL Quarterly* 9, 2, 173-183.

Krashen, S., J. Butler, R. Birnbaum, and J. Robertson. 1977. Two studies in language acquisition and language learning. Paper presented at the 1st Annual Second Language Research Forum, Los Angeles, February, 1977.

Loftus, G. 1975. A study of four teachers' lessons. In R. L. Allwright (ed.).

Long, M. H. 1976. Encouraging language acquisition by adults in a formal instructional setting. *ELT Documents*, (1976/3), British Council, London.

Lucas, E. 1975. Teachers' reacting moves following errors made by pupils in post-primary English-as-a-second language classes in Israel. MA thesis, Tel Aviv University.

Mackworth, N. H. 1950. Researches on the measurement of human performance. *Med. Res. Council Special Report*, No. 268.

Mehan, H. 1974. Accomplishing classroom lessons. In A. Cicourel *et al.* (eds.) *Language use and school performance*, New York, Academic Press, 76-142.

Moerk, E. 1976. Processes of language teaching and training in the interactions of mother-child dyads. *Child Development*.

Nelson, K. E., G. Carskaddon, and J. D. Bonvillian. 1973. Syntax acquisition: impact of experimental variation in adult verbal interaction with the child. *Child Development 44*, 497–504.

Plann, S. 1977. Second language acquisition in the Spanish immersion program: towards native-like proficiency or a classroom dialect? Paper presented at the Eleventh Annual TESOL Convention, Miami Beach, Florida.

Richards, D. R. 1974. On communicative efficiency and the treatment of error. Unpublished paper, University of Essex, Department of Language and Linguistics.

Rowe, M. B. 1969. Science, silence and sanctions. *Science and Children 6*, 6.

Slobin, D. 1971. *Psycholinguistics*. Glenview, Ill. Scott, Foresman.

Stokes, A. 1975. Errors and teacher-student interaction. In R. L. Allwright (ed.).

Zahorik, J. A. 1968. Classroom feedback behaviour of teachers. *Journal of Educational Research* 62, 4, 147–150.

Practice Effect in English Proficiency Testing

J. Donald Bowen

American University in Cairo
UCLA

At the American University in Cairo the 12 or so forms of the MTELP (Michigan Test of English Language Proficiency) have been used for some 18 years as the principal evidence to determine admission to the university and, when special training in English is indicated, to decide on promotion to the regular university program. It is important to know whether a "practice effect" on readministration of this test causes an inflation of scores.

Data gathered in the present study, in which 38 students took five forms of the MTELP, indicates a modest and statistically nonsignificant mean gain (about .7 point per administration), well within the standard error of measurement. Fluctuation of means, and more so of individual scores (averaging 10.6 points, and with 84% of the subjects showing a variance of over 5 points), and inconsistency in the direction of score change, suggest caution in interpreting student competence from test scores alone, with the recommendation that if possible other evidence be considered, especially in borderline cases.

English-curriculum educational institutions that admit as students nonnative speakers of English have an understandable interest in English-language proficiency tests. These tests help decide which candidates are linguistically prepared for enrollment, which need special language training, and perhaps which are so weak in English that the institution would not want to undertake a special program to remedy their deficiency. Typically the institution administers to applicants a standard proficiency test, such as the MTELP (Michigan Test of English Language Proficiency) or the ALIGU (American Language Institute of Georgetown University) or the CELT (Comprehensive English Language Test), or employs a similar locally set test, or asks applicants to provide scores from a comparable externally administered test, usually the TOEFL (Test of English as a Foreign Language).

In some cases the same or a similar examination is used in the admissions action and subsequently to help in the decision to promote students to full admission status after a period of special training in English-language skills. In any of these circumstances the institution will be very much concerned with the validity and reliability of the instrument used. Especially when retests are given to make promotion decisions, the reliability of the test is very important. Wrong decisions in the admissions/promotion process cause problems that are painful and difficult to correct; consequently should be avoided if at all possible.

The present study addresses a problem associated with testing to determine if promotion from a special English training program to full student status is justified. It involves practice affect, the possibility that there is a gain in measurement scores that can be attributed specifically to test experience, and that evaluation scores should therefore be adjusted upward to allow for this mechanical gain when students are retested. The study was conducted at the American University in Cairo, where an admissions battery of tests is offered to all candidates seeking enrollment. The battery consists of the MTELP, the MTAC, and a locally-set written composition test. The equated scores of each of these tests are combined to produce a percentage-scale score on which the decision to enroll largely depends. It should be pointed out that the American University in Cairo, usually referred to as AUC, is an American-Egyptian private school that uses English as the medium of instruction in all programs except Arabic Studies, but that the student body is over ninety per cent Egyptian and to an even higher percentage Arabic-speaking. Academic success at AUC is not possible without fluent control of English language skills—both written and spoken.

The Admissions Battery classifies applicants into three groups: 1) those who are permitted to enroll directly in the freshman year, 2) those invited to enroll for an intensive language course (25 hours a week) in the English Language Institute (ELI) for up to three terms, 3) those denied admission. The cut-off points on the scores from the Admissions Battery are:

82-100—direct admission
63- 81—enrollment in the ELI
60- 62—a borderline group who may be admitted to the ELI program if they can produce compensating high scores on the Egyptian Thanawiyya Amma (the national secondary school-leaving examination). The combined score must be 140, which means the candidate in the Admissions Battery twilight zone must produce the relatively high Thanawiyya Amma score of 78 to 80.[illegible]
0 - 59—rejected

It may be of some interest to know what percentages of applicants fall in these three groups. The relevant data for the latest group of applicants tested in September 1976 are shown in Table 1. As can be seen from these data, the number of borderline admissions is very low.

In addition to its use in determining enrollment, the Admissions Battery is readministered at the end of each term to all the enrollees in the ELI. By the same patterns of performance (modified by a new factor of teacher evaluation of overall proficiency), students are passed on to full admission in the university or are retained for further training in the ELI. If the minimum score of 82 is not achieved after three terms of training, the student is dropped from the university. A student may, then, take the Admissions Battery as many as

TABLE 1

Performance of AUC Applicants

Category			Number	Percentage
Direct admission to AUC			123	22.4
Eligible for training at ELI			254	46.4
Clear (63-81)	245	44.7		
Borderline (60-62)	9	1.6		
Rejected			155	28.3
Clear (0-59)	139	25.4		
Borderline (60-62)	16	2.9		
Dropouts (incomplete data)			16	2.9
Total			548	100.0

four times—one to gain assignment to the ELI and three times attempting to move out of the ELI to regular status as a university student.

Students do not retake the same form of the Michigan tests a second, third, or fourth time. There are some ten or twelve froms available, so they will always be given a form they haven't had. Understandably, there is keen interest in the test, and active student support can be counted on for any classroom activity that is seen as enhancing the probability of improved test performance. There is a corresponding diminution of interest in anything that is judged indifferent. A few years ago it was decided to include the judgment of the teachers in the promotion process, based on two factors, one academic and one administrative: 1) A teacher who has observed and evaluated a full term's work by the student is in a very strategic position to render judgment on his competence and potential. Carried out seriously on the basis of common and agreed criteria, teacher judgment improves the evaluation process. 2) The administrative consideration was to encourage the student to serious participation in class activities, rather than having him succumb to the temptation of believing that home study, possibly with assistance from a private tutor, would do more to promote the likelihood of passing the next administration of the Admissions Battery, or call it the Promotions Battery.

It is very much to the interest of the University to know a number of things about testing and specifically about the Michigan tests. One factor of considerable importance is test security. Every effort is made to insure the protection of the tests, but admittedly there is an enormous temptation for compromise. The problem is kept within bounds by the numerous forms available for administration and by carefully withholding until the last minute information on which form will be used in a particular administration. Still two forms have been withdrawn because of the likelihood of compromise.

More crucial to the testing program are two other factors: 1) Are the various forms of the test in fact equivalent as measures of proficiency, and 2) Is there a learning effect inherent in sitting for the examination?

For the first question AUC has accepted the assurances of the test makers that if the equation formulas they supply are applied to raw scores, the results will indeed be equivalent measures. The second problem has been more nagging. A considerable local feeling has developed reflecting the belief that retest scores are not comparable to original scores—that the student who initially tests high will be more likely to achieve academically than the student who retests up from his original test score.

The present study was designed to determine if there was any merit in the assumption that the practice afforded by taking a test could by itself result in increased scores on subsequent administrations. It was planned and arranged to give five tests, all of them forms of the Michigan Test of English Language Proficiency, to students in the secretarial studies program of the Division of Public Service of the American University in Cairo.

The order in which the forms were presented was staggered so that each form in one of the administration groups occupied each slot in the sequence of five. The groups, arbitrarily numbered, and the sequence of forms administered to each group are listed in Table 2.

TABLE 2

Administrations by Group

Group	Sequence					N
	1	2	3	4	5	
1	F	G	H	J	K	7
2	H	K	J	F	G	9
3	G	J	F	K	H	8
4	K	F	G	H	J	6
5	J	H	K	G	F	7
6	F	J	H	G	K	1
Total						38

It will be noted that the groups are of slightly different numbers. All groups started out with the same number of subjects, but it was not possible to predict who would finish the series and who would drop out. The balanced pattern shown was adopted as an alternative to a set of patterns established by reference to a chart of random numbers. This procedure had to be abandoned for the very practical reason that the random-numbers-established sequence frequently schedules the same test for administration to more than one group at a time, and the number of test forms available did not permit a duplication of this kind. The balanced pattern that was adopted seems satisfactory, and there is no reason to believe it was biased.

A confusion of test forms established an unplanned group 6 with one subject. He should have been in group 1, but produced a unique sequence when he did form J on the second administration.

The students in the sample were selected because they represented a range of competence and ability comparable to applicants for undergraduate admission in the regular program of the university. Also they were in a university program, but not regular undergraduates who would need to take the test to determine their own promotion.

It was hoped that 80 to 100 subjects would be induced to sit for the tests, each test taking a maximum of 75 minutes. The secretarial studies program enrolled 168 students in a two-year course who were contacted to find out how many would be willing to take the series of tests. Seventy-nine agreed to participated and filled out questionnaires requesting demographic data. To encourage participation and to provide a measure of motivation (to at least attempt to match the considerable motivation that regular applicants and students have to succeed at the tests), participants were first assured that taking the five examinations would be a significant educational experience that would benefit them in their own future, and in addition they were offered compensation at the rate of one piaster for each point they earned on each of the five tests—if they completed the series. This meant a total of £E5.000 could be earned by a maximum performance on the five tests. (This can be compared to the modest monthly salary of about £E30.000 for beginning teachers in Egyptian public schools.) The compensation offered for sitting for the five tests was not munificent, but was a reasonable extra dividend for participation.

It was hoped that the tests could be given in the space of 8 days to minimize the effect of outside influence, as the participants could be presumed to be learning English from the secretarial training program they were following. Sessions three times a week were scheduled and participants were informed that they could attend any five of the six sessions.

Plans do not always work out as they are originally devised. For the first administration of the test, only 56 students appeared. There was a heavy dropout after that, especially among the academically weaker students, when those who came out of curiosity decided the project was too onerous. They were granted program time for the tests, but the alternative, since classes were cancelled, was free time, which some students apparently value highly. The second session produced only 26 students, the third 32, the fourth and fifth 27 each, and the sixth 9. At that time only 22 students had taken the entire series of five forms of the test. This was disappointing, but an end-of-term schedule pattern made starting with a new group of students unfeasible. Salvaging the project seemed to depend on encouraging the students who had taken some tests to continue. Accordingly, small groups of two or three or four were tested as they could be assembled until a total of 38 sets of five administrations were accumulated. This required a period of just over ten weeks. A total of 38 sets of scores is only about half of the minimum first contemplated, but that figure was arbitrary. Such are the problems involved in experimenting with human subjects, especially when cooperation depends on good will. At least I am able to take advantage of an unplanned opportunity—to determine if the score movement (up or down) would be any different for subjects who took the examinations within a relatively short interval of time as compared to those who took them over a longer interval.

The Michigan Test of English Language Proficiency—the part of the Admissions Battery involved in this study—consists of three subtests: grammar (40 items), vocabulary (40 items), and reading comprehension (20 items—five for each of four reading passages). All items are in a multiple-choice format and were therefore scored completely objectively. The usual ELI AUC scoring procedure was followed, which provided generous rechecking to insure accuracy. The test sessions were carefully monitored to assure that no information passed among students during the test.

In addition to the five forms of the Michigan Test, each student also sat for a 19-minute taped integrative grammar test as an external validation for their performance on the Michigan Test. A further opportunity for validation is provided by course grades in English and translation for those who successfully finished the academic term.

Means for the five administrations of the Michigan Test are given in Table 3. The N in all cases is 38. There are five scores for each test, raw scores for the grammar, vocabulary, and reading subtests, and a raw score and an equated score for the total performance. The test scores can be sequenced by performance, from low to high, as shown in Table 4.

TABLE 3

MTELP Performance Means
by Sequence

Administration	GR	VOC	RD	TOTAL	EQUATED
First	20.89	16.32	7.00	44.21	62.21
Second	20.95	16.79	7.50	45.42	63.03
Third	21.92	16.92	8.26	47.11	64.76
Fourth	20.16	16.84	8.34	45.32	63.16
Fifth	21.76	17.03	8.63	47.72	65.24
Average	21.14	16.78	7.95	45.89	63.68

TABLE 4

Data Summary

	Lowest			Highest		Lowest Mean Score	Highest Mean Score	Interval	% of Variance
GR	4	1	2	5	3	20.16	21.92	1.76	8.7
VOC	1	2	4	3	5	16.32	17.03	.71	4.3
RD	1	2	3	4	5	7.00	8.63	1.63	23.2
TOT	1	4	2	3	5	44.21	47.42	3.21	7.2
EQUA	1	2	4	3	5	62.21	65.24	3.03	4.8

By examining summary data from the scores we can see there is a quite regular tendency toward gain for each administration, though the gain is consistant only in the case of the reading test. The grammar test skews the results, which carries over into the raw total and equated scores. Besides being consistent in gain scores, the reading test shows gains that are considerable, with an interval of almost 25%.

It is possibly relevant to note that in the format of the MTELP the reading test comes last, and the three subtests are not separately timed. Some examinees may not save an adequate amount of time for the reading test on the early administrations, or they may become tired and, inattentive near the end of the test before they gain specific test experience, but learn to pace themselves better on later administrations, thus producing better scores. It could be argued that self-pacing and coping with mental fatigue are precisely the kind of elements of "test wisdom" that practice effect is concerned with, and that therefore the reading test result might well be the best evidence of practice effect.

But the project concerns the entire MTELP, and the consistent improvement in reading scores is contradicted, mildly by the vocabulary test and substantially by the grammar test, so that the raw and equated total scores fail to show a consistent gain, with administration four out of place two slots in the raw total scores and one slot in the equated total scores. This ought to be enough evidence to serve as a warning that practice cannot be counted on to produce gain scores, but some critics might observe that not many program participants take the test a fourth time, so it might still be a consistent measure for most examinees.

But are the gains, aside from being of questionable consistency, sufficiently significant that we should recommend an adjustment of the scores on successive administrations? In order to determine if the variance in the sequenced tests was statistically significant, a t-test for related samples was applied to each two adjacent administrations and to administrations one and five. All failed to show significance, even at the .20 level, which very convincingly demonstrates that the small gain in mean scores (which figures to an average of .7 of a point per test) cannot be relied on to indicate the presence of any learning from the practice effect (indeed even when other, non-test-related learning might possibly have taken place). The statistical test confirms the earlier judgment that there is no significant learning from practice.

To see if the scores for the reading tests alone showed a gain that was statistically significant, the same t-test for related samples was applied to the first and fifth administration scores, where a maximum gain of mean scores has been noted. As in the case of the equated scores, no significance was indicated, even at the .20 level. It should perhaps be pointed out that the reading scores are raw, not equated, since the equation formula applies only to the total raw scores. Nonetheless, form comparisons were well mixed, and there is no reason to suspect bias on a scale that would change the significance calculations.

From the first to the fifth administration there is an interval of just three points on the equated scores, but as indicated earlier, this gain tends to fluctuate. Gains (and losses) between administrations are shown in Table 5. This sequence shows a characteristic feature of the analysis. The score movements, by administration, tend to fluctuate around a combined mean. If a score in one adjacent comparison rises somewhat, the next score is likely to fall. This will be shown later as a prominent feature of individual performance even more so

TABLE 5

Gains and Losses

1-2	2-3	3-4	4-5
+.82	+1.73	−1.60	+2.08

than when comparing means, where most score reversals tend to be smoothed off. Note that the gain between administrations three and five, not counting four, is only .48.

It would be nice if all our decisions could be applied to average or typical cases. Then the statistician could take over and solve our problems. But almost no students are average, and many individual cases perform at quite a distance from the means. How far should we trust "average" solutions based on performance means?

In the present study we can look at the movement of scores, up and down, to see if the change in scores is consistent for individuals. Thirty-eight individuals each took a sequence of five tests, which allows us to look at 152 movements. A tally that indicates how many moved up, down, or stayed in one location between each sequence of two administrations provides information on the relative consistency of performance by individual examinees. The key that interprets significance assumes that to be recognized, a movement must exceed two points, the approximate standard error of measurement, so:

+	up more than two points
0	no change more than two points
—	down more than two points

The tally of individual score movement is shown in Table 6.

TABLE 6

Individual Score Movements

1-2			2-3			3-4			4-5			Totals		
+	0	—	+	0	—	+	0	—	+	0	—	+	0	—
15	10	13	20	9	9	8	16	14	17	16	5	60	51	41

It can be noticed that the plusses and minuses fluctuate: from 1-2 to 2-3 the number of plusses goes up, from 2-3 to 3-4 it goes down, from 3-4 to 4-5 it goes up again. Not surprisingly the minuses show a converse pattern. On the whole there are more plusses than minuses, though the criteria for a gain on retesting might well suggest that the zeros and minuses should be compared to the plusses, in which case the non-plusses are more numerous than the plusses.

Adding the number of points involved in the movements up and down allows us to calculate the average movement. For the complete data, the number of points on all occasions where gain was registered is 462; the average number of points on all occasions where loss was registered is 354. The difference of

108 points, divided by 152 opportunities for comparison of adjacent scores, yields an average gain of .71 per test, the average gain figure cited earlier. (This figure ignores the zeros—i.e., fluctuations within two points.)

Since there were some retests after relatively longer intervals of time, the opportunity is available to make a comparison between score movements after a short interval versus a longer interval. The movements of scores was divided into two groups: 1) score changes where the retest was given within nine days, and 2) score changes where the retest was given after twelve or more days. (There were no intervals of ten or eleven days.) The advantage should lie with the retests after longer intervals, since the subjects were being instructed throughout the entire period of testing.

There were 130 short-interval changes and 22 long-interval changes, ranging from 12 to 59 days. For the short-interval reexaminations, gains of 410 points exceeded losses of 271 points, producing a difference of 139 points divided by 130 opportunities for change yielding an average gain of 1.07 points per retest. For the long-interval reexaminations, gains of 52 points were offset by losses of 81 points producing a net loss of 29 points, which divided by 22 opportunities for change yielded an average loss of −1.32 points per retest. (The overall average gain of .72 then lies between the 1.07 gain for short-interval retests and the −1.32 loss for long-interval retests.)

Why should this be so? Why should longer intervals between testing be associated with a loss of points earned in the tests? One hypothesis is that the practice effect exists briefly after a test is taken, but wears off in time. Another explanation is that there were too few cases in the long-interval group to base conclusions on. A third explanation, in consideration of the need to round up strays to get enough cases to make the study worthwhile, is that the not-so-serious students were the dilatory ones who waited to be coaxed, and they may not have taken the testing so seriously.

This explanation appears to be confirmed by course grades for performance in the Secretarial Training program. The mean score in the English language course by the short-interval students was 80.1 ($N = 20$); the mean for the long-interval students was 66.9 ($N = 14$). The conclusion, drawn from the present data, seems to be that practice effect, if it exists, is limited to serious and high-performance students, and then is held to an average of about one percentage point per retest.

The movement of score data allows one more observation, for the consistency of change in individuals. One can ask, does a student move consistently up (or down) when he is retested several times. A consistency tally was made for all subjects, this time counting any gain and any loss. The tally is based first on any sequence of three tests and then on any sequence of four. The three-test sequences and the tallies of score movement are shown in Table 7 (Two plusses indicate two consecutive score gains on retests, two minuses indicate two consecutive score losses; The heading "Other" indicates mixed performance of gains, losses, or no movement.)

TABLE 7

Consistency of Score Movements for Individuals

1-2-3			2-3-4			3-4-5		
+ +	− −	Other	+ +	− −	Other	+ +	− −	Other
6	3	29	6	3	29	5	2	31

Of a total of 114 opportunities only 25 show consistency (17 up and 8 down), which constitutes an inconsistency proportion of 78.1%.

Looking at sequences of three retests is even more unbalanced, as shown in Table 8. Of 76 opportunities, only one is consistent through three retests, for an inconsistency index of 98.7%.

TABLE 8

Sequences of Three Retests

1-2-3-4			2-3-4-5		
+ + +	− − −	Other	+ + +	− − −	Other
0	0	38	1	0	37

One can easily see that a change in scores—up or down—on a single retest is a poor prediction of another similar change on a further retest. In fact the chances are substantially better for a contradictory change, which suggests that individuals should perhaps *not* be judged too conclusively on the basis of test scores—at least if test sophistication is suspected as a factor in performance.

To further emphasize the inadvisability of depending too confidently on a single test to measure language competence, a tally was made of the range of variation by individual subjects tested. How much variance was there in test scores for the same subject on the five tests? They ranged from a low of two to a high of 25 points. But the average variation was surprisingly high: 10.6 points (with a standard deviation of 5.50 to show the substantial distribution of individuals, 67 per cent between 5 and 16 points and another 17 per cent above 16 points). Given the lack of consistency in score movements and the tendency to fluctuate above and below an individual subject's personal mean, this seems to offer conclusive evidence of the hazardous position of placing too much confidence in test scores. Only four individual subjects varied within the four points that would be allowed in a theoretical provision for a one-point gain adjustment per test administration, and two of those were in the high nineties, with no room to fluctuate. None of the four went consistently up.

Since the design of the study lends itself to a comparison of forms this comparison was made. Regardless of the sequence in which the tests were taken, the means for individual forms is as follows. Again N is 38 in all cases.

Since the alphabetical order of the form designations is irrelevant, the listing is based on performance, as shown in Table 9. Scores by Form, low to high, are shown in Table 10.

TABLE 9

MTELP Performance Means by Form

Form	GR	VOC	RD	TOTAL	EQUATED
G	20.61	15.63	7.61	43.84	62.76
H	21.76	16.05	7.42	45.21	63.92
F	21.39	17.11	7.18	45.16	64.89
K	21.18	17.26	8.55	47.18	63.34
J	20.74	17.84	8.97	47.55	63.47
Average (as cited earlier)	21.14	16.78	7.95	45.89	63.68

TABLE 10

Data Summary

	Lowest			Highest		Mean Score Lowest	Mean Score Highest	Interval	% of Variance
GR	G	J	K	F	H	20.61	21.76	1.15	5.5
VOC	G	H	F	K	J	15.63	17.84	2.21	14.1
RD	F	H	G	K	J	7.18	8.97	1.79	24.9
TOT	G	F	H	K	J	43.84	47.55	3.71	8.4
EQUA	G	K	J	H	F	62.76	64.89	2.13	3.3

Since sequence, being randomized, is not a factor, presumably the low-score tests are more difficult and the high-score tests are easier. There is most variance on the reading tests, considerable variance on the vocabulary test, and little on the grammar test. Apparently the equation formulas are working, since the variation for the equated scores is reduced to a manageable 3.3 per cent (about two-thirds of the 4.8 per cent of variance on the comparison by sequence). Still this is almost seven per cent if one must calculate the possible range of error on scores presented by an individual on two forms of the test.

The intervals and percentages of variation by sequence and by form, juxtaposed for comparison, are shown in Table 11.

It is interesting that in both sets of comparisons the reading section shows

TABLE 11

Variation by Sequence and by Form

	Sequence		Form	
	Interval	%	Interval	%
GR	1.76	8.7	1.15	5.5
VOC	.71	4.3	2.21	14.1
RD	1.63	23.2	1.79	24.9
TOT	3.21	7.2	3.71	8.4
EQUA	3.03	4.8	2.13	3.3

the most variance. Roughly: variance by form is as great as by sequence. Perhaps as indicated earlier this is because the test, while taken within a maximum time of 75 minutes, is not timed by section, and the reading section coming last may be slightly in terms of time allotment. Still almost all students finish and leave before the 75 minutes is up. Perhaps there is more guessing or more wild guessing. (On one occasion I observed a student marking choices on his answer sheet for the reading test without reference to the test booklet.)

Grammar skills seem to be measured fairly consistently by form, but vocabulary shows considerably more variation by form than by sequence. Possibly consistency is more elusive in the open-ended area of vocabulary than in the relatively more confined area of structure—at least in the tradition of what gets tested.

The central questions of most interest to AUC, and perhaps to other institutions that use the Michigan Test, are: 1) are the forms equivalent within a sufficiently narrow range so that equated scores can be relied on to identify competence levels, and 2) can readministrations of alternate forms be depended on for reliable measures of improved competence. The answer to both questions is a qualified yes. There is always a chance of error in measuring human competences, and there is a margin of error in the forms of the Michigan test. But if this error is small, we can live with it.

As always, it is at the cut-off points where injustice is most likely. I believe AUC has wisely chosen to rely on additional data in its admissions exam procedure for the "twilight" zone of the near-miss scores just below qualification for training in the ELI, using as further evidence the performance on the Thanawiyya Amma. One minor difficulty is that not *all* applicants present a Thanawiyya Amma score, and it is not always easy to establish comparable levels of performance on the test scores that they do present.

But in addition to the twilight zone between rejection and ELI there is another cut-off point, the "sunrise" area between ELI and direct admission. Possibly applicants with scores of 80-81 could similarly be considered for direct admission if other evidence established the probability that their Admissions Battery scores were on the low side of the range of possible measurement error. Possibly an auxiliary or "caboose" test could be considered for this decision.

One possibility would be to add a test like the Integrative Grammar Test to help make the decisions at the borderlines. Such a test should be used only as a procedure to resolve questions of interpretation just below the cutoff points, not applied to all applicants. Perhaps the test should be given only in borderline cases or should be scored only for their resolution.

The Integrative Grammar Test (IGT) is practical for this purpose; it takes 19 minutes to administer and could be given as a caboose to the aural comprehension test of the Admissions Battery. It has proved to be highly reliable (r = .968—calculated by the equivalent form method for 686 subjects) and valid (r = .871 with grammar subtest of The Michigan Test of English Language

Proficiency, .866 with the MTELP itself, and .812 with the entire Admissions Battery, calculated for 632 subjects.[1]

The IGT was administered to the subjects in the present practice effect project with a correlation of .829 between the mean equated total for all five administrations and the IGT, and .871 between the IGT and the grammar subtest of the Michigan exam, shown in Table 12. These coefficients are very high, particularly since a 19-minute IGT is being compared to over six hours of testing for the combined five forms of the MTELP.

TABLE 12

Correlations of MTELP and IGT

MTELP	Integrative Grammar Test		
	G1	G2	G3
GR	.8540	.8682	.8714
EQUA	.8269	.8131	.8294

If the IGT were used as an auxiliary measure to confirm borderline judgments, break points would have to be established. I would suggest that these be a little higher than the mean performances of the reference group that have taken the test in the past. This performance is shown in Table 13.

TABLE 13

Performance Data on the IGT

	Mean	SD	Range	N
Direct Admission	64.9	15.52	17-98	110
ELI Assignment	36.1	14.72	2-77	190
Rejected	12.6	10.13	0-54	332

The separation of the means for the direct admissions and the ELI students is 28.8 points which is almost as high as the combined standard deviations for the two groups. I would suggest that the crucial points be 65 and 40, so the pattern for assignment would be as shown in Table 14.

If such a plan were adopted, it should of course be monitored to provide assurance that it works adequately. If advisable, new forms of the present test should be produced. And research on other, hopefully improved auxiliary instruments, should be continued to be sure the best current ideas are being applied as a clear policy. The ideal auxiliary test, for borderline duty, needs to be simple, short, practical, reliable, and valid—i.e., a good test. But it does not

---

[1] IGT—an experimental grammar test under development at AUC for the past 2 yrs. Data from "Current Research on an Integrative Test of English Grammar," *RELC Journal* (Singapore) 7:2, December 1976, 30-37.

TABLE 14

Suggested Admissions/Promotion Testing Policy for AUC

	Score on Adm. Batt.	Score on IGT
Direct Admission		
Clear	82-100	
Borderline	80-81	65+
ELI Assignment		
Clear	63-81	
Borderline	60-62	40+

need to be as elaborate as the more sophisticated instrument that does the main job of classifying applicants for admission.

In summary, there seems to be no significant gain from practice effect, but the readministration of forms of the Michigan Test of English Language Proficiency appears to indicate that while there is no overriding evidence to suggest that the tests are not valid and reliable, there is enough variation in scores to indicate that results should probably be accepted only as approximations to a valid, accurate measure of competence. Perhaps this is true for most tests, and a little skepticism is a healthy antidote to blind faith in measurement instruments.

Though the MTELP is functioning satisfactorily in the Admissions/promotion Battery at the American University in Cairo, individual variation of scores on retests suggests caution in the interpretation of test scores. It seems advisable to continue to use other evidence of linguistic and academic preparation for judging admissibility to AUC, especially the applicants' Thanawiyya Amma scores, when these are available. It also seems advisable to utilize other tests that have proved in experimental use to be valid and reliable, especially to help interpret competence and determine action in borderline cases. Such a policy should not only improve selection procedures, but should be popular with applicants, who would interpret the additional tests as an effort to make a more fair decision.

Using Rater Judgments in the Evaluation of Writing Proficiency for Non-native Speakers of English

Karen A. Mullen

University of Iowa

An investigation has been made to determine the equivalence of composition ratings in assessing second-language writing proficiency. One hundred and seventeen non-native speakers of English were asked to write a composition under controlled conditions. The subjects were arbitrarily divided into eight groups. Five experienced ESL teachers were paired to form eight sets of raters. Each group of compositions was read by one pair of evaluators and rated on five scales of writing proficiency. A two-factor analysis of variance was performed to test for a significant difference between groups and a single-factor analysis of variance was performed to test for a significant difference between raters within a pair. Unbiased reliability coefficients were derived from the partitioned variance. A stepwise regression analysis was performed to determine the relative weight given to each scale in assessing overall proficiency and to determine if all four scales were necessary to best predict the overall scores. The results of this investigation show no significant difference among the ratings assigned to the eight groups of subjects and a significant difference between raters within a pair. The evidence suggests that some rater pairs are able to reach equivalence in their judgments, other pairs can produce parallel but reliable judgments, and some pairs produce non-equivalent, non-parallel judgments. The results of this study also show that although scale ratings correlate highly, all four scales will do better in predicting the overall score than any three, any two, or any single scale. Furthermore, the results indicate that the ratings on the vocabulary usage scale play the heaviest role in determining the overall quality of the composition and that ratings on compositional organization play the least.

In the field of second language testing, one of the skills commonly tested is that of writing ability, either objectively in the context of multiple choice questions or productively in the framework of a writing task. Objective writing tests, most notably of the type included in the TOEFL, have been constructed to meet the requirements of reliability and validity. Yet one of the major criticisms is that they do not allow one to see how a second-language learner organizes his thoughts on paper or applies his known vocabulary, nor does it indicate how well a learner uses in extended, unified prose the formal grammatical rules he has been taught chapter by chapter in his language texts. On the other hand, productive writing tests have been criticized for their failure to produce reliable measurements of writing skill. This unreliability has been attributed to two sources: the topics on which the learner is writing and the

321

judges who evaluate the result. These criticisms have been levelled against such tests given to native speakers of English. However, for non-native speakers, the case may be different, particularly if the criteria to be measured are those of sentence structure, vocabulary usage, fluency of writing, and coherence of ideas. If the purpose of having a learner of English write is to test his ability to put sentences together while appropriately selecting from his store of vocabulary and to apply within a set period of time the grammatical rules he knows, then the question of whether the judges of such writing can assess this and produce parallel assessments is an important one. The purpose of this paper is to report the results of a study designed to determine if experienced ESL teachers, working in pairs, can come to a mutual agreement concerning the writing proficiency of non-native speakers of English and to determine the reliability of such judgments. In addition, the question of whether different sets of judges rate differently is posed. Finally, the role each scale plays in the evaluation of overall writing proficiency is examined. Specifically, the hypotheses are as follows:

1. There is no significant difference between the ratings assigned by judge 1 and judge 2 to a group of subjects in each of the several scales of writing proficiency.
2. There is no significant difference between ratings assigned by one pair of judges to one group of speakers and by another pair of judges to another group of speakers from the same sample.
3. The information provided by the four scales together can predict the overall writing proficiency score significantly better than any three, any two, or any single scale.

Experimental Design

To test hypothesis one, a single-factor experimental design having repeated measures was chosen. The F-statistic based upon the mean sum of squares between judges divided by the mean sum of squares of the residual variance was computed to test the hypothesis of no significant difference between judges. Unbiased reliability coefficients were calculated based upon the number of subjects in the sample, the number of judges within the group, the mean square between subjects, and the mean square within subjects (Winer, p. 287). To test hypothesis two, a two-factor experimental design having repeated measures on one factor was chosen. The F-statistic based upon the mean square between groups divided by the mean square of subjects within groups was computed to test the hypothesis of no significant difference among groups. To test hypothesis three, tthe F-statistic based upon the increment in the sum of squares due to the addition of a scale-variable divided by the residual variance was calculated from a stepwise regression analysis.

Procedure

The judges were required to rate the subjects on five scales of writing proficiency: control over English structure, organization of material, appropriateness of vocabulary, quantity of writing, and overall writing proficiency. The scales were labelled vertically on a rating form. Each scale was presented in the form of a double horizontal line equally divided into five contiguous compartments labelled from left to right: poor, fair, good, above average, and excellent. The judges were instructed to put an X in the box best characterizing the learner's proficiency with regard to each of the five scales or to put an X on the line between boxes if the evidence warranted. A set of guidelines for deciding what level of proficiency to assign was explained to the judges before they read the writing and it was at hand for reference during the evaluation. For analysis, judgments were later converted to a numerical value of 1 = poor, 2 = between poor and fair, 3 = fair, 4 = between fair and good, 5 = good, 6 = between good and above average, 7 = above average, 8 = between above average and excellent, and 9 = excellent.

The subjects were given a composition booklet with instructions inside directing them to choose a topic from one of four choices, to plan their ideas for ten or fifteen minutes, to develop their ideas using details and examples, and to consider that their writing would be evaluated for grammar, vocabulary, paragraph organization, logical development, and quantity of writing. Subjects were allowed an hour for the task.

Judges

Five judges participated in this study. They were randomly paired together to form eight groups. All judges were graduate students in linguistics. They had completed courses in phonetics, syntactic and phonological analysis, and TESL methodology. They had all taught ESL for at least a year. All had been instructed on how to use the rating form and the guidelines, and they had participated in such composition evaluation before. None of the judges had had a previous acquaintance with the subjects whose compositions they read.

Subjects

The one hundred and seventeen subjects in this study had been referred to the University of Iowa Linguistics Department for a proficiency evaluation by either the foreign admissions officer, the foreign student advisor, or the student's academic advisor. Most of the subjects were new to the university. Most had been referred because their TOEFL scores were below 550. Some appeared for an evaluation because in the course of their few days on campus, the foreign student advisors had noted a lack of facility in English, although the TOEFL scores were not below 550. The purpose of the evaluation was to determine whether additional instruction in English and a reduced academic program might be recommended for the student.

Results

Table 1 shows the results of a two-factor analysis of variance having repeated measures on one factor. A separate analysis of variance and two F-statistics are reported for each scale. The F-statistic for a difference between groups

TABLE 1

An Analysis of Variance of Performance Scores of Eight Groups of Subjects Tested under Different Pairs of Judges on Five Scales of Writing Proficiency.

		Structure			Organization			Quantity		
Source of Variance	df	SS	MS	F	SS	MS	F	SS	MS	F
Between Groups	7	59.4564	8.4937	1.74	70.3841	10.0548	1.89	106.3545	15.1935	2.38*
Between Subjects within Groups (pooled)	109	533.4751	4.8942		579.2397	5.3141		697.1069	6.3954	
Between Judges within Groups (pooled)	8	7.1984	.0899	1.22	32.6357	4.0794	4.45**	30.2481	3.7810	5.29**
Residual (pooled)	109	79.8015	.7321		99.8641	.9161		78.7517	.7224	

		Vocabulary			Overall		
Source of Variance	df	SS	MS	F	SS	MS	F
Between Groups	7	30.2229	4.3175	.83	47.6729	6.8104	1.35
Between Subjects within Groups (pooled)	109	566.3923	5.1962		548.3868	5.0310	
Between Judges within Groups (pooled)	8	17.5796	2.1974	3.03**	14.2153	1.7769	3.29**
Residual (pooled)	109	78.9202	.7240		58.7845	.5393	

* significant at $p < .05$
** significant at $p < .01$

TABLE 2

Analyses of Variance on Performance Scores (one for each of eight pairs of subjects tested under a different pair of judges) on Five Scales of Writing Proficiency.

			Structure		Organization		Quantity		Vocabulary		Overall	
Gp	Source of Variance	df	MS	F	MS	F	MS	F	MS	F	MS	F
1	Between Subjects	22	7.519	15.40**	6.430	7.14**	5.179	4.69**	7.760	10.47**	6.027	10.30**
	Between Judges	1	1.760	3.61	3.673	4.08	8.695	7.87*	.195	.26	3.130	5.35*
	Residual	22	.488		.901		1.104		.741		.584	
2	Between Subjects	11	2.041	2.46	1.678	2.47	2.829	3.68*	1.212	1.70	1.136	3.12*
	Between Judges	1	.375	.45	15.141	22.18**	7.041	9.16*	8.166	11.47**	6.000	16.50**
	Residual	11	.829		.678		.768		.712		.363	
3	Between Subjects	18	6.076	8.25**	7.865	8.10**	9.769	12.70**	5.964	6.75**	6.976	12.17**
	Between Judges	1	.236	.32	.026	.03	5.157	6.71*	7.605	8.61**	1.684	2.94
	Residual	18	.736		.970		.769		.883		.573	
4	Between Subjects	19	3.993	6.98**	4.636	5.47**	4.972	22.10**	4.578	5.24**	4.067	14.65**
	Between Judges	1	.625	1.09	4.900	5.78*	.225	1.00	.400	.46	.225	.81
	Residual	19	.572		.847		.225		.873		.277	
5	Between Subjects	10	2.036	1.56	2.109	1.59	3.800	5.97**	2.381	3.20*	2.109	2.32
	Between Judges	1	2.909	2.22	.727	.55	1.636	2.57	.045	.06	.409	.45
	Residual	10	1.309		1.327		.636		.745		.909	
6	Between Subjects	9	8.561	14.14**	10.827	13.08**	15.672	20.01**	12.472	77.41**	12.672	253.00**
	Between Judges	1	.050	.08	.050	.06	.450	.57	.050	.31	.050	1.00
	Residual	9	.605		.827		.783		.161		.050	
7	Between Subjects	11	3.829	4.98**	5.257	4.89**	7.041	8.19**	3.439	4.05*	4.257	3.80*
	Between Judges	1	1.041	1.35	4.166	3.87*	2.041	2.37	.666	.79	2.666	2.38
	Residual	11	.768		1.075		.859		.848		1.121	
8	Between Subjects	9	2.311	2.36	1.472	2.05	2.800	5.04*	1.561	3.47*	2.050	4.15*
	Between Judges	1	.200	.20	4.050	5.65*	5.000	9.00*	.450	1.00	.050	.10
	Residual	9	.977		.716		.555		.450		.494	

* significant at p < .05   ** significant at p < .01

is not significant at the .05 level for four of the scales—structure, organization, vocabulary, and overall proficiency. It is not significant at the .01 level for any of the five scales. The F-statistic for a difference between judges within a pair is significant at the .01 level for every scale except structure.

In order to ascertain which pairs of judges might be responsible for producing significant differences, we must consult Table 2, showing the results of a single-factor analysis of variance having repeated measures for each of the five scales and for each of the eight groups of subjects. The F-statistic for a difference between judges' structure scores is not significant at either the .05 or the .01 level. It is significant at the .05 level for a difference between judges' organization scores for four pairs of judges (groups 2, 4, 7, and 8), between judges quantity scores for four pairs (groups 1, 2, 3, and 8), between vocabulary scores for two pairs (groups 2 and 3), and for overall scores for two pairs (groups 1 and 2). For six out of the eight groups, there is a significant difference in ratings on at least one scale. At the .01 level, there is a significant difference between judges organization scores for group 2, between vocabulary scores for groups 2 and 3, and for overall scores for group 2. The F-statistic shows no significant differences between judges' quantity scores at this level. Additionally, there is no significant difference between judges across rating scales for six out of the eight pairs. Pair 3 shows a significant difference on one scale (vocabulary). Pair 2 is the most deviant, showing a significant difference at the .01 level for three out of the five scales. This pair performed similarly on an identically designed experiment involving scales of speaking proficiency on a randomly selected group of subjects from the same pool.

The reliability coefficient is a measure of the degree to which an average of the raters' scores on a given scale is a good estimator of the subjects' true scores and as such indicates the percentage of the obtained variance in the distribution of scores which may be regarded as variance not attributable to errors of measurement. Squaring the reliability coefficient will indicate the accuracy of the prediction when the score assigned by one judge in a pair is used to predict the score given by the other judge. Table 3 shows the unbiased reliability coefficients for each pair of judges for all scales of writing proficiency. The coefficients range from a low of —.34 to a high of .99. Table 3 also shows the relationship between the F-statistic for a difference between judges and the reliability coefficients. Some cases show a significant difference and a high reliability (vocabulary scores in group 3, for example), indicating that the judges are not interpreting the scale in the same way. The scores they assigned are not equivalent but are closely parallel. There are also cases in which there is no significant difference between judges' scores but a rather low reliability (structure scores for group 5, for example). This indicates that the judges, though showing no overall difference in assigning scores, do not interpret the scale in any way consistent with one another across all subjects in the sample. Where there is no significant difference in judges' scores and a high reliability (group 6 for all scales), we may infer that both judges are interpreting the

TABLE 3

A Comparison of the F-statistics from Analyses of Variance (from table 2) and Reliability Coefficients of Judge Pairs.

		Structure		Organization		Quantity		Vocabulary		Overall	
Group	n	F	r_2	F	r_2	F	r_2	F	r_2	F	r_2
1	23	3.61	.92	4.08	.83	7.87*	.70	.26	.90	5.35*	.87
2	12	.45	.53	22.18**	−.34	9.16*	.45	11.47**	−.32	16.50**	.12
3	19	.32	.87	.03	.87	6.71*	.89	8.61**	.77	2.94	.90
4	20	1.09	.84	5.78*	.75	1.00	.95	.46	.79	.81	.92
5	11	2.22	.13	.55	.26	2.57	.77	.06	.65	.45	.50
6	10	.08	.92	.06	.91	.57	.94	.31	.98	1.00	.99
7	12	1.35	.75	3.87*	.70	2.37	.84	.79	.71	2.38	.65
8	10	.20	.51	5.65*	.10	9.00*	.55	1.00	.64	.10	.73

* significant at $p < .05$

** significant at $p < .01$

scales similarly and reach agreement in their evaluation of individual subjects. In cases where there is a significant difference between judges and a low reliability (as for pair 2 on all scales), one may conclude that the judges disagree and that the error of measurement on the scales for these raters is sufficient to reduce the extent to which an average of the two scores could be used as an estimate of the subject's true score. Reliability coefficients for four of the eight pairs of judges are within the acceptable limits of .70 or above on all scales. A fifth pair is within acceptable limits on four out of the five scales. Three sets of judges did not produce acceptable reliability coefficients on the five scales. The quantity scale appears to provide the most uniform reliability coefficients across all groups of subjects and raters.

As shown in Table 4, the correlation between scale ratings is high. However, a stepwise multiple regression analysis reveals that as the variables are added one by one to the equation for predicting the overall proficiency rating, each addition to the equation significantly reduces the amount of unpredicted variance. The variable accounting for the most variance is added first and subsequent variables are then included by the same principle. Table 5 shows these results. The beta coefficients in the equation indicate that if the vocabulary rating were increased by one unit and the other ratings remained constant, the expected change in the overall proficiency rating would be .36. If the quantity rating were increased by one (and the other ratings remained constant), the change in the overall rating would be .23. Parallel conclusions may be reached for the influence of an increase in the structure or organization rating on the overall score. A unitary increase in the vocabulary rating would produce the most change (by about a third) and such an increase in the organization rating would produce the least (by about a fifth). Unitary increases in the quantity or structure scores would change the overall rating by about a fourth. If all four ratings were increased by one unit, the change in the overall proficiency rating would be very slightly more than one unit as well. This indicates that although the scales are unequally weighted, they function together in such a way that they produce a unitary change in the overall score. In addition, the

TABLE 4

Correlations of Ratings on the Five Scales of Writing Proficiency.

Scale	2	3	4	5
1 Structure	.79	.67	.84	.89
2 Organization		.82	.82	.90
3 Quantity			.72	.83
4 Vocabulary				.92
5 Overall Proficiency				

TABLE 5

A Stepwise Regression Analysis with the Overall Proficiency Scale as the Dependent Variable.

Source	df	SS	MS	F	r^2	Beta	S.E.(B)	F(for B)
Regression	1	564.874	564.874	1257.85**	.844			
Vocabulary	1	564.874	564.874	1257.85**		.92	.025	1257.85**
Residual	232	104.185	.499					
Regression	2	606.665	303.332	1123.02**	.906			
Vocabulary	1	564.874	564.874			.66	.028	521.74**
Quantity	1	41.791	41.791	154.72**		.36	.024	154.72**
Residual	231	62.393	.270					
Regression	3	626.911	208.970	1140.33**	.937			
Vocabulary	1	564.874	564.874			.42	.032	164.65**
Quantity	1	41.791	41.791			.31	.020	168.53**
Structure	1	20.246	20.246	110.47**		.32	.030	110.47**
Residual	230	42.148	.183					
Regression	4	631.583	157.895	964.82**	.943			
Vocabulary	1	564.874	564.874			.36	.032	126.14**
Quantity	1	41.791	41.791			.23	.023	70.13**
Structure	1	20.246	20.246			.27	.030	84.59**
Organization	1	4.672	4.672	28.54**		.18	.032	28.54**
Residual	229	37.476	.163					

** $p < .01$

R-square coefficient indicates that the four variables account for .943 of the variance in the overall scores and that the relationship between the four scales and the overall scale is very nearly linear.

Conclusions

It is apparent from the results of this study that some pairs of judges achieve fairly high reliability and show no significant difference in scoring on all scales. It is also apparent that some pairs of judges are highly reliable in their rating although one in the pair is calibrating its judgments consistently higher than the other. It is also clear that some pairs of judges cannot produce reliable judgments. Moreover, each of the scales plays a role in the determination of the overall score. In answer to the hypotheses stated at the beginning, we may conclude that:

1. There is evidence to reject the claim that there is no significant difference between the ratings assigned by judges within a pair.
2. There is no evidence to reject the claim that there is no significant difference in ratings assigned by the eight pairs of judges.
3. There is no evidence to reject the claim that all four scales together can predict the overall writing proficiency score better than any three, or any two, or any single scale.

REFERENCES

Winer, B. J. 1971. *Statistical principles in experimental design*, 2nd ed. New York, McGraw-Hill.

APPENDIX A

COMPOSITION EVALUATION

Name ______________________________ Date ____________________

Evaluator ____________________________

	Poor	Fair	Good	Above Aver.	Excellent
Control over English structure					
Compositional Organization					
Quality of Writing					
Appropriateness of Vocabulary					

Overall Writing Proficiency					

APPENDIX B

GUIDELINES FOR EVALUATION OF COMPOSITIONS

Control over English Structure

Excellent:	Few, if any, noticeable errors of grammar or word order. Frequent use of complex sentences.
Very Good:	Occasional grammatical and/or word-errors. Some use of complex sentences.
Good:	Frequent grammar and word-order errors. General use of simple sentences.
Fair:	Many errors in grammar make comprehension difficult. Use of short basic sentences.
oor:	Severe errors in grammar and word order. No apparent knowledge of English.

Compositional Organization

Excellent:	Well developed introduction which engages concern of the reader. Use of in-

ternal divisions and transitions. Substantial paragraphs to develop ideas. Conclusion suggests larger significance of central idea.

Very Good: Obvious inclusion of an introduction, though not smoothly developed. Division of central idea into smaller parts, though paragraphs are lean on detail. Conclusion restates the central idea.

Good: Intent to develop central idea is evidenced, but only a few points are mentioned. The introduction or conclusion is very simply stated or may be missing. Occasional wandering from topic.

Fair: Limited organization. Thoughts are written down as they come to mind. No introduction or conclusion.

Poor: No organization. No focus. No development. No major consideration of topic.

Quality of Writing

Excellent: Writing is an easy task. Quantity seems to be no problem.

Very Good: Reasonable quantity for the hour. Writing flows without much hesitation.

Good: Enough writing to develop the topic somewhat. Evidence of having stopped writing at times.

Fair: Much time spent struggling with the task of putting down thoughts on paper.

Poor: Very little writing during the hour-long assignment.

Appropriateness of Vocabulary

Excellent: Precise and accurate word-choice. Obvious knowledge of idioms. Aware of word connotations. No translation from native language apparent. May have attempted a metaphoric use of words.

Very Good: Occasional misuse of idioms, but little difficulty in choosing appropriate forms of words. Uses synonyms to avoid repetition. Some vocabulary problems may be due to translations.

Good: Use of the most frequently occurring words in English. Does not use synonyms to avoid repetition. Some inappropriate word-choices. Uses circumlocutions or rephrasing when the right word is not available.

Fair: Depends upon a very small vocabulary to convey thoughts. Repetition of words is frequent. Appears to be translating. Great difficulty in choosing appropriate word forms.

Poor: Vocabulary is extremely limited.